

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 6, 2001, 10:53:40 ; Search time 16.05 Seconds
(without alignments)
1775.033 Million cell updates/sec

Title: US-08-878-801-2

Perfect score: 1960
Sequence: 1 MARSLTWRCPCWCLTEDEKA.....YKDYVDYLARYLDEINIL 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Maximum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first: 45 summaries

Database: PIR_68:*

1: p1r1:***
2: p1r2:***
3: p1r3:***
4: p1r4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1960	100.0	374	2 A41096	GTP-binding regula
2	1705	87.0	374	2 B41534	GTP-binding protei
3	1057	53.9	359	2 S43699	GTP-binding regula
4	1046	53.4	353	2 B40891	GTP-binding protei
5	1042	53.3	359	1 RGM60	GTP-binding regula
6	1038	53.2	359	1 S71963	GTP-binding protei
7	1035	53.0	359	1 RGM511	GTP-binding regula
8	1035	52.8	359	1 S30359	GTP-binding regula
9	1031	52.6	359	1 RGM67	GTP-binding regula
10	1026.5	52.4	359	1 S45700	GTP-binding regula
11	1017.5	51.9	354	2 S33309	GTP-binding regula
12	1014.5	51.8	360	2 JN0115	GTP-binding regula
13	1013	51.7	355	2 A40891	GTP-binding protei
14	1013	51.7	355	2 A41534	GTP-binding protei
15	995	50.8	353	2 S34347	GTP-binding regula
16	974	49.7	355	2 T15288	hypothetical prote
17	787.5	40.2	353	2 S71965	GTP-binding regula
18	766.5	39.1	353	2 T50482	G protein alpha ch
19	762	38.9	354	1 RGF02	GTP-binding regula
20	760	38.8	353	1 S25493	GTP-binding regula
21	756.5	38.6	354	1 RGF01	GTP-binding regula
22	755	38.5	354	1 S27013	GTP-binding regula
23	752.5	38.4	352	2 B41095	GTP-binding regula
24	750.5	38.3	354	1 T24154	hypothetical prote
25	748.5	38.2	354	1 S27014	GTP-binding regula
26	748.5	38.2	354	1 T50237	GTP-binding regula
27	746.5	38.1	354	2 S24352	gustducin - rat
28	746	38.1	377	2 I57490	guanine nucleotide

ALIGNMENTS

30	745.5	38.0	354	1 RGH011	GTP-binding regula
31	745.5	38.0	354	1 RGH011	GTP-binding regula
32	745.5	38.0	354	1 RGH011	GTP-binding regula
33	744.5	38.0	354	2 A61035	GTP-binding regula
34	744.5	38.0	354	2 S40508	GTP-binding regula
35	742.5	37.9	354	1 RGH02	GTP-binding regula
36	741.5	37.8	354	1 S28157	GTP-binding regula
37	740.5	37.8	354	2 T19476	hypothetical prote
38	740.5	37.8	354	1 RGH013	GTP-binding regula
39	740.5	37.8	354	1 RGH02	GTP-binding regula
40	740.5	37.8	354	1 RGH02	GTP-binding regula
41	739.5	37.7	354	1 RGH02	GTP-binding regula
42	739.5	37.7	355	2 A61031	GTP-binding regula
43	738.5	37.7	354	1 RGH013	GTP-binding regula
44	738	37.7	355	1 RGH012	GTP-binding regula
45	737.5	37.6	354	2 S28159	GTP-binding regula

RESULT 1

A41096
GTP-binding regulatory protein alpha-16 chain - human
C:Species: Homo sapiens (man)
C:Date: 20-Mar-1992 #sequence_rev1sion 20-Mar-1992 #text_change 02-Feb-2001
C:Accession: A41096
R:Amatuda III, T.T.; Steele, D.A.; Slepak, V.Z.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 88, 5587-5591, 1991
A:Title: Galphai6, a G protein alpha subunit specifically expressed in hematopoietic
A:Reference number: A41096; M01D:91288509
A:Accession: A41096
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <MAS>
A:Cross-references: GB:M63904; NID:g182891; PIDN:AAA35860.1; PID:g182892
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; nucleotide binding; P-loop
F:49-56/Region: nucleotide-binding motif A (P-loop)
F:277-280/Region: GTP-binding NKXD motif

Query Match	100.0%	Score 1960;	DB 2;	Length 374;
Best Local Similarity	100.0%	Pred. No. 9.2e-147;		
Matches 374;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MARSLTWRCPCWCLTEDEKAARVDQINRILLEOKKODRGELKLLIGGESCKSTFIK	60	
DB	1	MARSLTWRCPCWCLTEDEKAARVDQINRILLEOKKODRGELKLLIGGESCKSTFIK	60	
QY	61	QMRITGAGSEERKGFRLVYONIFVSRAMIEAMERQIPPSRESKHHSLVNSOD	120	
DB	61	QMRITGAGSEERKGFRLVYONIFVSRAMIEAMERQIPPSRESKHHSLVNSOD	120	
QY	121	PKVTTPEKRYAAMOMLMDAGIRACYERREPHILDSAVYVSHLERTTEETVPTAQ	180	
DB	121	PKVTTPEKRYAAMOMLMDAGIRACYERREPHILDSAVYVSHLERTTEETVPTAQ	180	
QY	181	DVIRSRMPTTGINEYCFVQKTMRLVDVGOKSERKKWICENVTALITLASLSEYDQ	240	
DB	181	DVIRSRMPTTGINEYCFVQKTMRLVDVGOKSERKKWICENVTALITLASLSEYDQ	240	
QY	241	CLENNQENMRKESLAFGLITLPMFKSVTLITLTKTILTEKIPTSILATYFPEOG	300	
DB	241	CLENNQENMRKESLAFGLITLPMFKSVTLITLTKTILTEKIPTSILATYFPEOG	300	
QY	301	PKDAAEAKRFTLDMYTRMTGCVDPGEGSKGARSRLFSHYTCATDTONIRKVFQVR	360	
DB	301	PKDAAEAKRFTLDMYTRMTGCVDPGEGSKGARSRLFSHYTCATDTONIRKVFQVR	360	
QY	361	DSVLARYLDEINIL 374		
DB	361	DSVLARYLDEINIL 374		



Sequence Comparison B

source

source

ORIGIN

Perce

Align

2

920 CACGAGAGGANTACACAGCAGTCCCTGGAGAGIACAACGAGGAAACCG 969
251 MellysglSerLeuAlaLeuphegLYTHrILLeuGluleProTrrPh 267
970 ATGAGAGAGAGCCCTCGCATTTGTTGGACATATCTGGAACTACCTGGT 1019
267 elySerThSerValleLeupheLeuAnlysrThAspIleLeuGlug 284
1020 CAAGAAGACATCCGTCACTCTTTTCCACACAAACCGACATCTCGAGG 1069
284 lulySleProThSerHisLeuAlaThrTyrrheProSerPheGingly 300
1070 AGAAATCCCCACCTCCACCTGCTACCACTATTTCCCGAGTTCCAGGGC 1119
301 ProLySglAspAlaGluAlaAlaAlaYsArgPheILeLeuAspMetTyrrh 317
1120 CCAAGAGAGATGCTGAGGACGACCAAGAGGTTATCTGAGCATGTACAC 1169
317 rArgMetTyrrhGlyCysValAspGlyProGlyuglySerlySlyGlyYA 334
1170 GAGGATATACACCGGGGCGCTGAGACGGCCCGGAGGACAGAAAGAGGG 1219
334 lArgSerArgArgLeupheSerHisTyrrhThcYsAlaThrAspTrpGln 350
1220 CAGCATCCGACGCTTTTACGCACTATACAGATGTCCACAGACACACAG 1269
351 AsnIleAlaGlyValPheLyAspValArgAspSerValLeuAlaArgTy 367
1270 AACATCGCAGAGCTCTTCAAGACAGTCGCGGACCTCGTCTGCGCCGCTA 1319
367 rLeuAspGluIleAsnLeuLeu 374
1320 CCGAGCAGATCAACCTCTG 1341

seq_name: gb_pr10: H0MGAL6

LOCUS documentation_block: 2060 bp. mRNA PRI 27-APR-1993
DEFINITION Human G-alpha 16 protein mRNA, complete cds.
ACCESSION M63904
VERSION M63904.1 GI:182891
KEYWORDS G-alpha 16 protein.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
Amatruda,T.F.III., Steele,D.B., Slepak,V.Z. and Simon,M.I.
G-alpha16, a G protein alpha subunit specifically expressed in
hematopoietic cells
Proc. Natl. Acad. Sci. U.S.A. 88, 5587-5591 (1991)
91288509

JOURNAL
MEDLINE
FEATURES
SOURCE

1. 2060
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HL-60"
1. 2060
/product="G-alpha-16 protein"
220. 1344
/codon_start=1
/product="G-alpha-16 protein"
/protein_id="AAA5860.1"
/db_xref="GI:182892"
/translation="MARSLTWRCQPMWCLFEDERKAARVDOELNRIILLKOKKDRGELK
LLTLPGESGKSTFKKMRILIGAGYSEERKRPILVYONIVSRNMIEMERLOI
PFPSRKHAAHSLVMSQDPADYITFERKAAANQMWRAGICACERRERFHLDSA
VYVLSHERITTEGVEYTPADYITRNRPTTGINIEYFVQKTNLRIVDVGCGSERKK
WIRFDEKILPTSHLATYPPS:OGPKDENDEENRKRBSIALGFTLELPPKRSVLLF
NKIDLELEKIPTHSLATYPPS:OGPKDENDEENRKRBSIALGFTLELPPKRSVLLF
RSRRLESHTYCAWDTONIRKRV:KDVDSVYLARIYDEINLL."

BASE COUNT 415 a 631 c 640 g 374 t

[illegible]

RESULT 2

B41534

GTP-binding protein alpha-15 chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 02-Feb-2001

C:Accession: B41534

R:Wikie, T.M.; Scherle, P.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.

Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053, 1991

A:Title: Characterization of G-protein alpha subunits in the G-q class: expression in mu

A:Reference number: A41534; MUID:92052208

A:Accession: B41534

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <WIL>

A:Cross-references: GB:M80632; NID:g193570; PIDN:AAA37713.1; PID:g193571

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop

F:49-56/Region: nucleotide-binding motif A (P-loop)

F:277-280/Region: GTP-binding NKXD motif

Query Match 87.0% Score 1705; DB 2; Length 374;
 Best Local Similarity 84.8% Pred. No. 1.1e-126;
 Matches 317; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

QY 1 MARSLTWRCPCWCLTEDEKAAARVDQINRILLLEOKKQDRGELKLLLLGPGESGKSTFIK 60
 DB 1 MARSLTWRCPCWCLTEDEKAAARVDQINRILLLEOKKQDRGELKLLLLGPGESGKSTFIK 60
 QY 61 QMRTIHGAGYSEERKGFRLVYONIFVSMRAMIEMERLQIPSRPESKHHASLVMSQD 120
 DB 61 QMRTIHGAGYSEERKGFRLVYONIFVSMRAMIEMERLQIPSRPESKHHASLVMSQD 120
 QY 121 PYKVTFEKYYAAMQWLMDRAGIRACYYRRRHHLLDSAVYLSHLRTEREGYVPTAQ 180
 DB 121 PYKVTFEKYYAAMQWLMDRAGIRACYYRRRHHLLDSAVYLSHLRTEREGYVPTAQ 180
 QY 181 DVLRSMPPTGNEYCFSVQKTLNRIVDVGGQSKERKKWHCFENFVIALIYLSLSYDQ 240
 DB 181 DVLRSMPPTGNEYCFSVQKTLNRIVDVGGQSKERKKWHCFENFVIALIYLSLSYDQ 240
 QY 241 CLEENQENRMKESALFGLTLELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSG 300
 DB 241 CLEENQENRMKESALFGLTLELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSG 300
 QY 301 PKODAAKRFILDMYRMYTCVGDGPEGKKGARSRLFSHYTCATDTONIRKVKFDVR 360
 DB 301 PRDAAAKRFILDMYRMYTCVGDGPEGKKGARSRLFSHYTCATDTONIRKVKFDVR 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 DSVLARYLDEINLL 374

RESULT 3

S45699

GTP-binding regulatory protein alpha chain q - African clawed frog

N:Alternate names: G-alpha-q protein

C:Species: Xenopus laevis (African clawed frog)

C:Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001

C:Accession: S45699

R:Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battey, J.F.

FEBS Lett. 348, 83-92, 1994

A:Title: Neurokinin B receptor, expressed in Xenopus laevis oocytes, selectively couples

A:Reference number: S45699; MUID:94298961

A:Accession: S45699

A:Molecule type: mRNA

A:Residues: 1-359 <SRA>

A:Cross-references: GB:U10502; NID:g505689; PIDN:AAA52189.1; PID:g505690

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop; signal transduction

F:46-53/Region: nucleotide-binding motif A (P-loop)

F:274-277/Region: GTP-binding NKXD motif

F:52/Binding site: GTP (Lys) #Status Predicted

F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 53.9% Score 1057; DB 2; Length 359;

Best Local Similarity 56.4% Pred. No. 1e-75;

Matches 204; Conservative 61; Mismatches 85; Indels 12; Gaps 2;

QY 13 CLTDEKAAARVDQINRILLLEOKKQDRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 10 CLSEAEKARRINDEIERQLRRDKDRARRELLLLLTGTGSGKSTFIKQRIIHGAGYSD 69
 QY 73 EERKGFRLVYONIFVSMRAMIEMERLQIPSRPESKHHASLVMSQDPYKVTTFEKRYA 132
 DB 70 EDRKGFRLVYONIFVSMRAMIEMERLQIPSRPESKHHASLVMSQDPYKVTTFEKRYA 129
 QY 133 AAMQWLMDRAGIRACYYRRRHHLLDSAVYLSHLRTEREGYVPTAQDVLRSRMPPTGI 192
 DB 130 DAIKYLWMDPGIQECYDRREYQLSDSTKYVINVDRIATQGYLPTQDVLRSRMPPTGI 189
 QY 193 NEYCFSVQKTLNRIVDVGGQSKERKKWHCFENFVIALIYLSLSYDQCLLENQENRMK 252
 DB 190 IEYFPDLQSVIFRMDVGGQSKERKKWHCFENFVIALIYLSLSYDQCLLENQENRMK 249
 QY 253 ESLALFGTILLELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSGPKODAAAKRFI 312
 DB 250 ESKALFTIITYPWFQNSVILFNKTDILEEKIPTSHLATYFPFSGPKODAAAKRFI 309
 QY 313 LDYRMYTCVGDGPEGKKGARSRLFSHYTCATDTONIRKVKFDVRSVLARYLDEIN 372
 DB 310 LKWEVDL-----NPDSK-----IYSHFTCATDTONIRKVKFDVRSVLARYLDEIN 357
 QY 373 LL 374
 DB 358 LV 359

RESULT 4

B40891

GTP-binding protein GL2 alpha chain - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001

C:Accession: B40891

R:Nakamura, F.; Ogata, K.; Shiozaki, K.; Kameyama, K.; Ohara, T.; Nukada, T

J. Biol. Chem. 266, 12676-12681, 1991

A:Title: Identification of two novel GTP-b. nding protein alpha-subunits that lack app

A:Reference number: A40891; MUID:91286303

A:Accession: B40891

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-353 <NAK>

A:Cross-references: GB:D90336; NID:g217567; PIDN:BAAL4350.1; PID:g217568

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop

F:40-47/Region: nucleotide-binding motif A (P-loop)

F:150-152/Region: GTP-binding SAK/L motif

F:268-271/Region: GTP-binding NKXD motif

Query Match 53.4% Score 1046; DB 2; Length 353;

Best Local Similarity 56.1% Pred. No. 7.3e-75;

Matches 203; Conservative 64; Mismatches 83; Indels 12; Gaps 2;

QY 13 CLTDEKAAARVDQINRILLLEOKKQDRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 4 CLSDEVESKRINAEIERQLRRDKDRARRELLLLLTGTGSGKSTFIKQRIIHGAGYSE 63
 QY 73 EERKGFRLVYONIFVSMRAMIEMERLQIPSRPESKHHASLVMSQDPYKVTTFEKRYA 132
 DB 64 EDRKGFRLVYONIFVSMRAMIEMERLQIPSRPESKHHASLVMSQDPYKVTTFEKRYA 123

QY 133 AAMOWLRDAGIRACVRRRERPHLLDSAVYYLSHLERITEEGYVPTAQDVLSRMPTTGI 192
DB 124 SAIKLWNDPGIOECYDRREYQLSDSAKYIUTVDRIATSGYLPTQODVLVRVPPTTGI 183
QY 193 NEYCFSVQKTNLRIYDVVGQSKERRKHICFENVTALIIYLASISEYDOCLENNNOENRMK 252
DB 184 IEYPFDLENIIFPMVDVGQSRERRKHICFENNVSIMELVALSEYDQVLVESDNENRME 243
QY 253 ESIALFGTTILELPWFKSTSVILFNKNTDILEEKIPTSHLATYFPSPQGPKQDAEAAKRFI 312
DB 244 ESKALERTIVTYPFQGNSSVILLFNKKOLLEDKILHSLHLDVFPEPDGPDQDAQAAREFI 303
QY 313 LDWYTRMYTCGVDPGSGKGARSRLFSHYCATDTONIRKFVKDVRDSVLARYIDEIN 372
DB 304 LKMFDVL-----NPDSDK-----IYSHFTCATDTENIRFVFAAKDTILNLNKEYN 351
QY 373 LL 374
DB 352 LV 353

RESULT 5
RGMSQ
N; G-protein binding regulatory protein Gq alpha chain - mouse
P; Alternate names: guanine nucleotide binding protein Gq alpha chain; heterotrimeric G-G
C; Species: Mus musculus (house mouse)
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: A38414
R; Strathmann, M.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990
A; Title: G protein diversity: a distinct class of alpha subunits is present in vertebrate
A; Reference number: A38414; PMID: 1067657
A; Accession: A38414
A; Molecule type: mRNA
A; Residues: 1-359 <SPR>
A; Cross-references: GB:M55412; NID:g193501; PIDN:AAA63306.1; PID:g193502
C; Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay
signals. The beta and gamma chains, required for GTPase activity, appear to be common to all
types. It is specific for each type of G protein.
C; Superfamily: GTP-binding regulatory protein Gs alpha chain
C; Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction
F: 46-53/Region: nucleotide-binding motif A (P-loop)
F: 274-277/Region: GTP-binding NKXD motif
F: 52/Binding site: GTP (Lys) #status predicted
F: 102/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match	53.3%	Score 1044	DB 1	Length 359
Best Local Similarity	55.5%	Pred. No. 1.le-74		
Matches 201	Conservative 62	Mismatches 87	Indels 12	Gaps 2
QY	13	CLTEDEKAAARVDQENIRLLLEOKKODRGELKLLLLPGESGKSTFIKOMRIIHGAGYSE	72	
DB	10	CLUSEAKAARRINDEIVRVRDKDRARELKLKLLLGCGESGKSTFIKOMRIIHGSGYD	69	
QY	73	EBRKGFPLVYQNIQFVSNRAMIEMERLQIPSRPESKHHASLVMSQDPYKVTTEKRYA	132	
DB	70	EDKRGFTKLVIQNIETAMQAMIRAMDTLPIKYEYHNKAHAOLVREVVDYKVSAPENPVV	129	
QY	133	AMQWLWRDAGTRACVERREPHLLDSAVYLYLSHLERITEEGVPTAQDVLKSRMPTTGI	192	
- DB	130	DAIKSLWNDPGIOEQCYDRRSQOLSOTIIYYLLNDLDRVADPSYLPQQDVLVRVPTTGI	189	
QY	193	NEVCFSVQKTNLRIVDVGOKSEKRWIKHCFENVIALIYIASLSEYDQCLEENQENRMK	252	
DB	190	IEYFPDLQSVIFRMVDVGGORSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME	249	
QY	253	ESIALFGTILELPPWFKSTSVILEFLNKTDIDILEEKIPTSHLATYFPSPGPKQDAEAAKRFI	312	
DB	250	ESKALPRTIITYPWFQNSVSSVILFLNKKDLLEEKIMYSHLVDYFPPEYDGPORDQAARFI	309	
QY	313	LDMYTRMYTCGVDPGPEGSKKANSRRLFSHYTCATDTQNIIRKVKFDVRDSVLVARI	372	

Db 310 LKMFVDL-----NPDSDK-----IIYSHFTCATDTFENIRFVAAVKDTTLQLNLKEYN 357

QY 373 LL 374
I: 375
Db 358 LV 359

RESULT 6
S71963
GTP-binding protein alpha-q - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 02-Feb-2001
C:Accession: S71963
R:Johnson, G.J.; Leis, L.A.; Dunlop, P.C.
Biochem. J. 318, 1023-1031, 1996
A:Title: Specificity of G-alpha(q) and G-alpha(11) gene expression in platelets and e
A:Reference number: S71963; MUID:96433124
A:Accession: S71963
A:Molecule type: mRNA
A:Residues: 1-359 <JOH>
A:Cross-references: EMBL:L76256; NID:q1478071; PIDN:AAB39498.1; PID:q1478072
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; nucleotide binding; P-loop; signal transduction
F:46-53/Region: nucleotide-binding motif A (P-loop)
E:274-277/Region: GTP-binding NKXD motif

	Query Match	53.2%	Score 1042;	DB 2;	Length 359;
	Best Local Similarity	55.8%	Pred. No. 1.5e-74;		
	Matches 201;	Conservative 60;	Mismatches 87;	Indels 12;	Gaps 2;
QY 13	CLTEDEKAAARVDQELNRLLEQKKODRGELKLLGLPGESGKSTFKOMRIIHGAGYSE 72				
Db 10	CLUSEAKEARRINDEIERQLRRDKRDARRELKLLLTGTGESGKSTFKOMRIIHGSGYSD 69				
QY 73	EPKRGFERPLVYQNIQVSMRAMIEAMERLOIPESRPESKHHASLVMSQDPYKVTTFEKRYA 132				
Db 70	EDKRGFTKLIVYQNIQTAMQAMIRAMDTLKIYKYEHNKAHAQLRVDDVEKVSAPFVY 129				
QY 133	AAOMWLWDAGIRACVYRRERREFLLDSAVYVLSHLERITEEGVYPTAODVLRSMPTGI 192				
Db 130	DAIKSLWDPGIDECYDRRREYQLSDSTKYKLLNDLRVADPAYLPTQDDVLRVRPTGI 189				
QY 193	NEYCFVSQVNTLRIVDYGQKSERKKWIHCFENVIALIYVLSASIEYDQCLEENNOENRMK 252				
Db 190	IETFPDLSQVIFRMVDVGGGSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249				
QY 253	ESLALFTGIIELPWFKSTSVILELNKTDILEEKIPTSHLATYFPFSGFGQDAAAKRFI 312				
Db 250	ESKALFTIITYPWFQNSVILELNKDKLLEKIMYSHLVDYFPPEYDGPQDAQAAREFI 309				
QY 313	LDMYTRMYTCVDGPEGSKGASRRFLSHYTCATDTQNTKRVKDKVRDSVLARYLDEN 372				
Db 310	LKMFVDL-----NPDSDK-----IIYSHFTCATDTFENIRFVAAVKDTTLQLNLKEYN 357				

RESULT 7
RGS11
GTP-binding regulatory protein G11 alpha chain - mouse
N:Alternate names: guanine nucleotide binding protein G11 alpha chain; heterotrimeric
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C:Accession: B38414
R:Strathmann, M.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990
A:Title: G protein diversity: a distinct class of alpha subunits is present in verte
A:Reference number: A38414; MUID:91067657
A:Accession: B38414
A:Molecule type: mRNA
A:Residues: 1-359 <STR>
A:Cross-references: GB:M55411; NID:q193499; PIDN:AAA63305.1; PID:q193500
C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that re
ains. The beta and gamma chains, required for GTPase activity, appear to be common to

ase; it is specific for each type of G protein.
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction
 F:46-53/Region: nucleotide-binding motif A (P-loop)
 F:274-277/Region: GTP-binding NKXD motif
 F:52/Binding site: GTP (Lys) #status predicted
 F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 53.0%; Score 1038; DB 1; Length 359;
 Best Local Similarity 55.5%; Pred. No. 3.2e-74;
 Matches 201; Conservative 65; Mismatches 84; Indels 12; Gaps 2;

QY 13 CLTEDEKAARVDQENIRILLLEOKKODRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 10 CLSDEVKESKRINAEIKQLRRDKRDARRELKLLLLGTGESGKSTFIKQRIIHGAGYSE 69
 QY 73 ERKGRPLVYONIFVSMRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKKRYA 132
 DB 70 EDKRGFTKLIVYONIFTAMQAMIRAMETLILYKYEONKANALLIREVDVEKVTTFEHOYV 129
 QY 133 AAMQWLWDAGIRACRYERREHLLDSAVYVLSHLERITEEGYVPTAODVLRSMPTTGI 192
 DB 130 NAIKTLWSDPGVQECYDRREFFQLSDSAKYVLTVDRIATVGYLPTQODVLRVPTTGI 189
 QY 193 NEYCSVQKTNLRIVDVGQSGKSRKWKHCFFENVIALYLSLEYDOCLEENNOENRMK 252
 DB 190 IEYFDLENILFRMVDVGGQSRERKWKHCFFENVTISIMFLVALSEYDQVLVESDNENRME 249
 QY 253 ESLALFGTILELPWFKSTSVILFNKTDILEKIPTSHLATYFPSPQPKQDAEAAKRFI 312
 DB 250 ESKALFTIITYPWFQNSVILFNKDLLEKILYSHLVDFYFDFGPDQDAQAAREFI 309
 QY 313 LDMYTRMTGCVGDPGSGKSGARSRLFSHYTCATDTONIRKVKDVRSDVLAARYLDEIN 372
 DB 310 LKMEVDL-----NPDSK-----IYSHFTCATDTENIRVFVAARKDTILQNLKEYN 357
 QY 373 LL 374
 DB 358 LV 359

RESULT 8

S30359
 N:GTP-binding regulatory protein Gs alpha chain, phospholipase C-activating - turkey
 C:Species: Meleagris gallopavo (common turkey)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Feb-2001
 C:Accession: S30359; S30360
 R:Naurice, D.H.; Waldo, G.L.; Morris, A.J.; Nicholas, R.A.; Harden, T.K.
 Biochem. J. 290, 765-770, 1993
 A:Title: Identification of Galpha(11) as the phospholipase C-activating G-protein of tur
 A:Reference number: S30359; MUID:93207527
 A:Accession: S30359
 A:Molecule type: mRNA
 A:Residues: 1-359 <MAW>
 A:Cross-references: GB:X73072; NID:g312254; PIDN:CAA51530.1; PID:g312255
 A:Experimental source: blood
 A:Accession: S30360
 A:Molecule type: protein
 A:Residues: 78-92;121-132;158-180;253-256;307-312;339-345;355-359 <MAW>
 A:Experimental source: erythrocytes
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:46-53/Region: nucleotide-binding motif A (P-loop)
 F:156-158/Region: GTP-binding NKXD motif
 F:274-277/Region: GTP-binding NKXD motif
 F:52/Binding site: GTP (Lys) #status predicted
 F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 52.8%; Score 1035; DB 2; Length 359;
 Best Local Similarity 55.5%; Pred. No. 5.5e-74;

Matches 201; Conservative 65; Mismatches 84; Indels 12; Gaps 2;
 QY 13 CLTEDEKAARVDQENIRILLLEOKKODRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 10 CLSDEVKESKRINAEIKQLRRDKRDARRELKLLLLGTGESGKSTFIKQRIIHGAGYSE 69
 QY 73 ERKGRPLVYONIFVSMRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKKRYA 132
 DB 70 EDKRGFTKLIVYONIFTAMQAMIRAMETLILYKYEONKANALLIREVDVEKVTTFEHOYV 129
 QY 133 AAMQWLWDAGIRACRYERREHLLDSAVYVLSHLERITEEGYVPTAODVLRSMPTTGI 192
 DB 130 NAIKTLWSDPGVQECYDRREFFQLSDSAKYVLTVDRIATVGYLPTQODVLRVPTTGI 189
 QY 193 NEYCSVQKTNLRIVDVGQSGKSRKWKHCFFENVIALYLSLEYDOCLEENNOENRMK 252
 DB 190 IEYFDLENILFRMVDVGGQSRERKWKHCFFENVTISIMFLVALSEYDQVLVESDNENRME 249
 QY 253 ESLALFGTILELPWFKSTSVILFNKTDILEKIPTSHLATYFPSPQPKQDAEAAKRFI 312
 DB 250 ESKALFTIITYPWFQNSVILFNKDLLEKILYSHLVDFYFDFGPDQDAQAAREFI 309
 QY 313 LDMYTRMTGCVGDPGSGKSGARSRLFSHYTCATDTONIRKVKDVRSDVLAARYLDEIN 372
 DB 310 LKMEVDL-----NPDSK-----IYSHFTCATDTENIRVFVAARKDTILQNLKEYN 357
 QY 373 LL 374
 DB 358 LV 359

RESULT 9

RGHUGY
 N:GTP-binding regulatory protein Gy alpha chain - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
 C:Accession: A39394
 R:Jiang, W.; Pandey, S.; Tran, V.T.; Fong, H.K.W.
 Proc. Natl. Acad. Sci. U.S.A. 88, 3907-3911, 1991
 A:Title: Guanine nucleotide-binding regulatory proteins in retinal pigment epithelial
 A:Reference number: A39394; MUID:91219481
 A:Accession: A39394
 A:Molecule type: mRNA
 A:Residues: 1-359 <GJA>
 A:Cross-references: GB:M69013; NID:g183690; PIDN:AAA58624.1; PID:g183691
 C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that re
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to
 rise; it is specific for each type of G protein.
 C:Genetics:
 A:Gene: GDB:GN11
 A:Cross-references: GDB:132587; OMIM:139313
 A:Map position: 19p13.3-19p13.3
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transductio
 F:46-53/Region: nucleotide-binding motif A (P-loop)
 F:274-277/Region: GTP-binding NKXD motif
 F:52/Binding site: GTP (Lys) #status predicted
 F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 52.6%; Score 1031; DB 1; Length 359;
 Best Local Similarity 55.2%; Pred. No. 1.1e-73;
 Matches 200; Conservative 66; Mismatches 84; Indels 12; Gaps 2;

QY 13 CLTEDEKAARVDQENIRILLLEOKKODRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 10 CLSDEVKESKRINAEIKQLRRDKRDARRELKLLLLGTGESGKSTFIKQRIIHGAGYSE 69
 QY 73 ERKGRPLVYONIFVSMRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKKRYA 132
 DB 70 EDKRGFTKLIVYONIFTAMQAMIRAMETLILYKYEONKANALLIREVDVEKVTTFEHOYV 129

[illegible]

RESULT 10
S45700
G-alpha-11 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C:Accession: S45700
R:Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battay, J.F.
FEBS Lett. 348, 89-92, 1994
A:Title: Neuromedin B receptor, expressed in Xenopus laevis oocytes, selectively couples
A:Reference number: S45699; MUID:94298961
A:Accession: S45700
A:Molecule type: mRNA
A:Residues: 1-339 <SHA>
A:Cross-references: GB:U010494; NID:g505687; PIDN:AAA52188.1; PID:g505688
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; nucleotide binding; P-loop
F:46-53/Region: nucleotide-binding motif A (P-loop)
F:274-277/Region: GTP-binding NKXD motif

Query Match	52.4%	Score	1026.5	DB 2	Length	359			
Best Local Similarity	54.0%	Pred. No.	2.6e-73						
Matches	202	Conservative	65	Mismatches	92	Indels	15	Gaps	3

1	MARSLTWRCPCWCLTDEKAAARVDQENRILLQKQDRGKELKLLILGGESGKSTFIK	60
1	CLSEVVEKSRINAEKEQLRRDKDSRRELKLLILGTGSGKSTFIK	57
1	MTLOSTWAC---	11
61	QMRIHGAGVSEERKGFRLPYQNTFYFSMRAMIEAMERLOIPESRPSKHHASLVNSOD	120
58	QMRIHGSGVSEEDKGFKLWLFQNTAMQSWIRAMETLKILYEQNKANAAQVVREDD	117
121	PYKVTTTFEKRYYAAAWLMDAGIRACYERRREFHLLDSAVYILSHLERITEEGYVPTAQ	180
118	VEKVCTFEPQYVNAIKNLWSDPGICEYDRRREYQLSDTKYYLTFDVRISKPGCLPTQQ	177
181	DVLSRMPPTGNEYCSFVQKTNLRIVDVGQGSERKWHCFENVIALIYSLASLSEYDQ	240
178	DVLSRVRPPTGTGIEYFDELNIIRFMDVVGQGSERKWHCFENVTSIMFLVALSEYDQ	237
241	CLEENNOENRMKESIALFGTILPWFKSTSVILFUNKYTDILBEKIPTSHLATVPSPFQG	300
238	VLVESONEHNMESKALFTIITYPWFQNSVILFLNKKDLLEDIMYSHLVDFVPFEDG	297
301	PKQDAEAAKRFILDMYTRMYTCGVDPGEGSKGARSRLFSHVTCATDTONIRKVKFDVR	360
298	PQDAATAAREFILKMFVDL-----NPDSDK-----IYSHFTCATDFTENIREVEAAVK	345
361	DSVLARYLDEINLL	374
346	DTILQHNLKEYNLV	359

RESULT 11

S33309

GGRP-binding regulatory protein Gq alpha chain - northern European squid

C:Species: Loligo forbesi (northern European squid)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001

C:Accession: S33309

R:Ryba, N.J.P.; Findlay, J.B.C.; Reid, J.D.

B:Biochem. J. 292, 333-341, 1993

A:Title: The molecular cloning of the squid (Loligo forbesi) visual Gq-alpha subunit

A:Reference number: S33309; MUID:93277493

A:Accession: S33309

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-354 <RYB>

A:Cross-references: EMBL:L10289

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; p-loop

F:40-47/Region: nucleotide-binding motif A (p-loop)

F:150-152/Region: GTP-binding SAK/L motif

F:269-272/Region: GTP-binding NKXD motif

Query Match	51.9%	Score	1017.5	DB 2	Length	354	3
Best Local Similarity	55.1%	Pred. No.	1.3e-72				
Matches	200	Conservative	57	Mismatches	93	Indels	13
Gaps							
13	CLTDEKAAARVDQEI	NRILLEOKQDRGELKLLILGPGESKSTFIKQMRIIHGAGYSE	72				
4	CLSEEAKEQRI	NEIEKQLRRDKDARRELKLLILGTGESKSTFIKQMRIIHGAGYSE	63				
73	BERKGFRPLVYIV	FSVMRAMIEAMERLQIPFSRPSKHHASLVMSQDPYKVYTTREKRYA	132				
64	EDRKGFEKIVIT	ONFSAIQTLLIAMAETLSLEYKDPNSNEHAELNSIDADSADIFEDGHV	123				
133	AAAMQLWRDAGI	TRACYERRRPFHLLDSAAVYVLSHLRITEEGVVPTAQDVLRSRMTPTGI	192				
124	TAIKGCWTD	DPGMOECYDRREYQLTDSKYYLLDDVERIHEPGVIPTQLILRVRVPTGI	183				
193	NEVCFSVQKTNL	RIVDVGQSKSERKWTCHCPENVALIYLASISEYDOCL-ENNNOENRM	251				
184	IEYFPDLYSI	IFRMVDVGQSERSEKWTCHCPENVTSTINFALVSEYDQVLVESDNEENRM	243				
252	KESLALPGTILE	LHPFKSTSVTLFNKNTDILLEKIPTSHLATYPPFQGPQKQDAEAKRF	311				
244	EESKALFRTIIT	VPWFQNSVTLFNKKDLLEEIKMTSHLADYPPDYGDKCDYEAAREF	303				
312	ILDMYTRMYT	CGVDYEGPESKKGAARSLPESHVYTCATDTONIRKVFQDVROSVLARYLDEI	371				

QY 372 NLL 374
 11:
 ppb 352 NLV 354

RESULT. 12

UNIONJ0115

GRTP-binding regulatory protein dqg alpha chain - fruit fly (*Drosophila melanogaster*)

Ni:Alternate names: dqg protein

C:Species: *Drosophila melanogaster*

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001

C:Accession: JN0115

R:Lee, Y.J.; Dobbs, M.B.; Verardi, M.L.; Hyde, D.R.

Neuron 5, 889-898, 1990

A:Title: dqg: a *Drosophila* gene encoding a visual system-specific G alpha molecule.

A:Reference number: JN0115; MUID:91097801

A:Accession: JN0115

A:Molecule type: DNA

A:Residues: 1-360 <LEE>

C:Genetics:

A:Gene: dqg

A:Cross-references: FlyBase:FBgn0004435

[illegible]

RESULT 13
 A40891
 Splicing protein GUL alpha chain - bovine
 Spliced from proteinus taurus (cattle)
 Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001
 Accession: A40891
 Nakamura, F., Ogata, K., Shiozaki, K.; Kameyama, K.; Ohara, K.; Haga, T.; Nukada, T.
 Biol. Chem. 266:12676-12681, 1991
 Title: Identification of two novel GTP-binding protein alpha-subunits that lack apparent
 Reference number: A40891; MUID:91286303
 Accession: A40891
 Status: preliminary
 Molecule type: mRNA
 Residues: 1-355 <NAK>
 Cross-references: GB:D90335; NID:q217565; PIDN:BAAL4349.1; PID:q217566
 Superfamily: GTP-binding regulatory proteins in alpha chain
 Keywords: GTP binding; nucleotide binding; P-loop
 42-49/Region: nucleotide-binding motif A (P-loop)
 152-154/Region: GTP-binding SAK/L motif
 170-273/Region: GTP-binding NKXD motif

```

Query Match      51.7%; Score 1013; DB 2; Length 355;
Local Similarity 54.4%; Pred. No. 2.9e-72;
Matches 139; Conservative 62; Mismatches 91; Indels 14; Gaps 2;

9 CPGWCTEEDKKAARVDQENIRILLLEOKKQDRGEUKLLLLGPGSGKSTFTKQMRIHCA 68
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4 CC--CLSAEKESQRTSASIEIQLQRDRDKDAKREKULLLLGTGSGKSTFTKQMRIHGS 61

```

QY	69	GYSEERKGFRLPVYQNI	EFVSNRA	IEAMER:QIPFSR	PESKHHASLW	SDPYKV	TTFE	128		
		:	:	:	:	:	:			
Db	62	GYSDERKGFLLVYQNI	FTAMQMI	RAMDTI:KIQYV	CEQNKENAQL	REVEVDK	VSTLS	121		
		:	:	:	:	:	:			
QY	129	KRYARAMQWLDAGI	RACRYER	RERFHL	LDLSAVYYL	SHLERIT	EEGYVPTAQD	VLRSMP	188	
		:	:	:	:	:	:			
Db	122	RQVEAIKQWQD	PGIQE	RRREYQ	LSDSIKYLD	IDIRIAMP	AFVPTQOD	VLRVVP	181	
		:	:	:	:	:	:			
QY	189	TTGINEYCF	SQVKTNL	RYDVGG	QKSERK	KWJHGFEN	VIALIYUAS	SEYDQCLEENNQE	248	
		:	:	:	:	:	:			
Db	182	TTGIIETYP	FDLENI	IFRMVD	VGGQSR	SRKWIHGF	SVTSILF	VLUSEYDQVLAEC	241	
		:	:	:	:	:	:			
QY	249	NRMKESLAF	GTILEL	PEWFK	STSVIL	FLNKTIDLEEK	IPITSHL	ATYPPSQGPKQDAEA	308	
		:	:	:	:	:	:			
Db	242	NRMESKALF	TIITYP	WFLNSS	VILFLN	KKDLLEEKIM	SHLISF	PTGTPQDVKAA	301	
		:	:	:	:	:	:			
QY	309	KRFTIOWMY	TRMYTC	VDG	PEGSKG	KGARSRLF	SHYTCAT	DTONIRKVFQD	VRSLARYL	368
		:	:	:	:	:	:			
Db	302	RDFILKLYQ	-----	-----	-----	-----	-----	-----	-----	361
		:	:	:	:	:	:			
QY	369	DEINLL	374	-----	-----	-----	-----	-----	-----	349
		:	:	:	:	:	:			
QY	350	REFNLV	355	-----	-----	-----	-----	-----	-----	355
		:	:	:	:	:	:			

RESULT 14
 41534
 TP-binding protein alpha-14 chain - mouse
 ;Species: Mus musculus (house mouse)
 ;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 02-Feb-2001
 Accession: A41534; E33033
 Wilkie, T.M.; Scherle, P.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.
 Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053, 1991
 Title: Characterization of G-protein alpha subunits in the G-q class: expression in
 Reference number: A41534; MUID:92052208
 Accession: A41534
 Molecule type: mRNA
 Residues: 1-355 <WIL>
 Cross-references: GB:M80631; NID:g193568; PIDN:AAA83222.1; PID:g193569
 Strathmann, M.; Wilkie, T.M.; Simon, M.I.
 oc. Natl. Acad. Sci. U.S.A. 86, 7407-7409 1989
 Title: Diversity of the G-protein family: sequences from five additional alpha-sub
 Reference number: A33833; MUID:90017488
 Accession: E33833
 Molecule type: mRNA
 Residues: 217-267 <STR>
 Cross-references: GB:M57616; NID:g193380; PIDN:AAA63304.1; PID:g193381; GB:M26739
 Superfamily: GTP-binding regulatory protein Gs alpha chain
 Keywords: GTP binding; nucleotide binding; P-loop
 42-45/Region: nucleotide-binding motif A (P-loop)
 152-154/Region: GTP-binding SAK/L motif
 270-273/Region: GTP-binding NKXD motif

Query Match	51.7%	Score 1013;	DB 2;	Length 355;
Local Similarity	54.4%;	Prod. No. 2,9e-72;		
Matches 199;	Conservative 61;	Mismatches 92;	Indels 14;	Gaps 2;
9	CCPCLCTEDEKAAARVDOENIRLLEQKQKODRGEJAKLLLLGPGESGSKSPFTKQRIINGA	68		
4	CC--CLSAEEKESORISAEIHRVRRDKDARREIKLLLLGTGSGSKSTFIKQRIILINGS	61		
69	GYSEBERKGPRLVYQNIYFVSMAHMEAMERQIIVPSRPESKHHASLVMSODPQKVITFE	128		
62	GYSDDRKGFTKLYQNIFTAMQAINRAMDTLRICYMCQNKENAOIIRREVVDKVTALS	121		
129	KRYAAMQWLWRDAGIRACVYRRREFHLLDSAVYIYLSHLERTTPEGVYPTAADYVLRSRMP	198		
122	RDQVAATIKQLMDPGTQECYDRRREYQSDSAKYITLTDIERIAMPSPFVPTQDDVLRVRVP	181		
189	TTGNIYCEFSVOKNLRIVDVGGQSKSRKKWTHCFENVIALIYLSLASEYDQCLEENNOE	248		
182	TTGIIIEYFPDENIIFRVDVGGQSKSRKKWTHCFESVTISFIYLSLASEYDQVLAECNDE	241		

us-08-878-801-2.rpr

Thu Sep 6 15:16:57 2001

QY 249 NRMKESLALFCTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPFOGPKQDAEAA 308
 Db 242 NMEEESKALFRTIITYPFLNSSVILFNKKDLLEEKIMYSHLSYFPEYTGPKQDVKAA 301
 QY 309 KRFIIDMYTRMYTGCVDGPEGSKGARSRLFSHYTCATDTONIRKVKFDVRDVSILARYL 368
 Db 302 RDFILKLYQ-----DQNPDKREKVIYSHFTCATDTENIRFVFAAVKDTILOLNL 349
 QY 369 DEINLL 374
 Db 350 REFNLV 355

RESULT 15

S34347
 GTP-binding regulatory protein Gq alpha chain - great pond snail
 C:Species: *Lymnaea stagnalis* (great pond snail)
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001
 C;Accession: S65461; S34347
 R;Knol, J.C.; Ramnatsingh, S.; van Kesteren, E.R.; van Minnen, J.; Planta, R.J.; van Hee
 Eur. J. Biochem. 230, 193-199, 1995
 A:Title: Cloning of a molluscan G protein alpha subunit of the Gq class which is express
 A:Reference number: S65461; MUID:95324523
 A:Accession: S65461
 A:Molecule type: mRNA
 A:Residues: 1-353 <RNO>
 A:Cross-references: EMBL:23106; NID:g312629; PIDN:CAA80653.1; PID:g312630
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction
 F:40-48/Region: nucleotide-binding motif A (P-loop)
 F:150-152/Region: GTP-binding SAK/L motif
 F:268-271/Region: GTP-binding NKXD motif
 F:46/Binding site: GTP (Lys) #status predicted
 F:177/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 50.8%; Score 995; DB 2; Length 353;
 Best Local Similarity 53.3%; Pred. No. 7.5e-71;
 Matches 193; Conservative 64; Mismatches 93; Indels 12; Gaps 2;
 QY 13 CLTEDEKAAARVDQINRILLLEOKKQDRGELKLLILGPGESGKSTFIKQRIIHGAGYSE 72
 Db 4 CIPDELKEQKRNQETIEROLKDRKDRRELKLLLTGTGSGSKSTFIKQRIIHGAGYSD 63
 QY 73 EERKGRPLVYQNIYFVSMRAMTEAMERLQIPSPRESKHASLVMSQDPYKVTTFEKRYA 132
 Db 64 EDRSHIKIYQNIYFAMHAMIRAMDTLNIQYINPANRENGNMIRQIDYETVTTFDKPCV 123
 QY 133 AAMQWLWRDAGIRACRYERRERFHLDSAVYLYSLHLERITEGYVPTAQDVLRSRMTTGI 192
 Db 124 DAIISLWDDGIQECYDRRREYQLTDSAKYILDSVERISQDYLTPLQDILRVVRVPTGI 183
 QY 193 NEYCFSVQKTNLRIVDVGQSKERKWIHCFFENVIALIYASLSEYDQCLENNOENRMK 252
 Db 184 IEYFDDLSIIFRMVDVGQSRERRKWIHCFFENVTSIMFVALSEYDQVVLVESDNENRME 243
 QY 253 ESLALFCTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPFOGPKQDAEAAKRFI 312
 Db 244 ESKALFRTIITYPWFQNSSVILFNKKDLLEEKIMYSHLSYFPEYTGPKQDAEAAKRFI 303
 QY 313 LDNRYTRMYTGCVDGPEGSKGARSRLFSHYTCATDTONIRKVKFDVRDVSILARYLDEIN 372
 Db 304 LKMFVEL-----NPDPK-----IYSHFTCATDTENIRFVFAAVKDTILOLNLKEYN 351
 QY 373 LL 374
 Db 352 LV 353

Search completed: September 6, 2001, 10:54:04
 Job time: 24 sec

Thu Sep 6 15:16:57 2001

us-08-878-801-2.rpr

Page 8

Thu Sep 6 15:16:56 2001

us-08-878-801-2.frag

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 6, 2001, 10:53:40 ; Search time 20.43 seconds

(without alignments)
1109.808 Million cell updates/sec

Title: US-08-878-801-2

Perfect score: 1960

Sequence: 1 MARSLSWRCPCWCLTEDEKA.....VFKDVRDVLARYLDELINLL 374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_0601:*
- 1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT:*
 - 2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT:*
 - 3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT:*
 - 4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT:*
 - 5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT:*
 - 6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT:*
 - 7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT:*
 - 8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT:*
 - 9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT:*
 - 10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT:*
 - 11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT:*
 - 12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT:*
 - 13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT:*
 - 14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT:*
 - 15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT:*
 - 16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT:*
 - 17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT:*
 - 18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT:*
 - 19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT:*
 - 20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT:*
 - 21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT:*
 - 22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1960	100.0	374	20	AAV49121
2	1960	100.0	374	21	AAV99841
3	1947	99.3	374	19	AAW42434
4	1711	87.3	374	21	AAV93970
5	1705	87.0	374	20	AAV49120
6	1701	86.8	374	19	AAW42435
7	1037	52.9	1276	20	AAV49127
8	1037	52.9	1394	20	AAV49129
9	1037	52.9	1397	20	AAV49134
10	1037	52.9	1418	20	AAV49131
11	1035.5	52.8	1323	20	AAV49133

12	1033.5	52.7	1303	20	AAV49132
13	1033	52.7	359	20	AAV52705
14	1033	52.7	359	20	AAV29789
15	1029	52.5	359	20	AAV49125
16	1013	51.7	355	21	AAV15026
17	746.5	38.1	354	19	AAW74713
18	746.5	38.1	354	21	AAV59514
19	746	38.1	377	20	AAV43317
20	745.5	38.0	354	21	AAV85290
21	742.5	37.9	354	14	AAV42424
22	740	37.8	355	21	AAV85149
23	738.5	37.7	354	21	AAV94207
24	731.5	37.3	350	14	AAV42426
25	731.5	37.3	350	21	AAV59536
26	725.5	37.0	354	15	AAV58924
27	711.5	36.3	354	14	AAV42425
28	711.5	36.3	354	21	AAV59535
29	707	36.1	381	21	AAV67285
30	692.5	35.3	354	15	AAV58914
31	654	33.4	394	19	AAW61187
32	651	33.2	380	21	AAV23382
33	647	33.0	394	19	AAW61186
34	646	33.0	394	19	AAW61183
35	643	32.8	394	17	AAV4559
36	640	32.7	394	19	AAW61188
37	639	32.6	394	19	AAW61180
38	639	32.6	394	19	AAW61184
39	637.5	32.5	470	20	AAV02227
40	637	32.5	394	19	AAW61182
41	637	32.5	394	19	AAW61185
42	636	32.4	394	19	AAW61179
43	633	32.3	389	19	AAW44948
44	633	32.3	394	19	AAW61181
45	622.5	31.8	472	20	AAV02224

ALIGNMENTS

RESULT 1

AAV49121

ID AAV49121 standard; Protein: 374 AA.

AC AAV49121;

DT 07-JAN-2000 (first entry)

DE G protein alpha 16.

XX G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.

OS Homo sapiens.

PN WO9951641-A1.

PD 14-OCT-1999.

PF 02-APR-1999; 99WO-US07333.

PR 03-APR-1998; 98US-0080671.

PA (NPSF-) NPS PHARM INC.

PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;

PI Simin RT;

DR WPI; 1999-610995/52.

XX

PT New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease -
 PS Claim 32; Fig 4; 255pp; English.
 XX
 CC The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a CaR
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABAR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
 CC linker. (1), and recombinant chimeric receptors (CR) without the GP
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used;
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of CaR and mGluR
 CC domains allows presentation of GABAR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 XX
 XX Sequence 374 AA;

Query Match 100.0%; Score 1960; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.6e-189;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MARSLTWRCPCWCLTEDEKAAARVDQENIRILLLEQKKQDRGELKLLLLGPGESGKSTFK 60
 DB 1 marsltwrcpcwcltedekaaarvdqenirillleqkkqdrgeklilllpgesgstfk 60
 QY 61 QMRIIHGAGYSEERKGFPLVYQNIIVSMRAMIEAMERLQIPFSPESKHHASLVMSQD 120
 DB 61 qmriihgagyseerkgfplvyqniivsmramieamerlqipfspeskhhaslvmseqd 120
 QY 121 PYKVTTFEKRYAAAMQWLWRDAGIRACYERRREFHLLDSAVYVLSHLERITEGYVPTAQ 180
 DB 121 pykvttfekryaaamqwlwrdagiracyerrrefhlldsavvylshleriteegyvypta 180
 QY 181 DVLSRMPPTGINEYCFVQKTNLRIVDVGGQKSERKKWHCFENVIALIYLASISEYDQ 240
 DB 181 dvlsrmpptgineycfvqktnlrivdvvggkserkkwhcfenvialiyilasisey 240
 QY 241 CLEENNOENRMKESLALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPSQG 300
 DB 241 cleennqenrmkeslalfgtilelpwfkstsvilfnktdileekiptshlatypsfg 300
 QY 301 PQQDAEAAKRFILDMYTRMYTCVDPGEGSKKGARSRLFSHYTCATDTQNIKRVFKDVR 360
 DB 301 pqqdaeaakrfildmytrmytcvdpgegskkgarsrlfshytcatdtqnirkvfkdv 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 dsvlaryldeinll 374

RESULT 2

AAAY99841
 ID AAY99841 standard; Protein: 374 AA.
 AC AAY99841;
 XX
 DT 08-SEP-2000 (first entry)
 XX Human G-alpha-16 protein.
 DE

XX
 KW Human; G-alpha-16; G protein; cytostatic; hyperproliferative disorder;
 KW cancer; inflammation; infection; antisense inhibition.
 XX Homo sapiens.
 XX WO200032817-A1.
 XX 08-JUN-2000.
 XX 25-AUG-1999; 99WO-US19613.
 XX 03-DEC-1998; 98US-0205143.
 XX (ISIS-) ISIS PHARM INC.
 XX Cowser LM;
 XX WPI: 2000-412354/35.
 XX N-PSDB; AAA48751.
 XX A new antisense compound for inhibiting the expression of human
 PT G-alpha-16 and treating, preventing or delaying infections,
 PT inflammation or hyperproliferative disorders such as cancer -
 XX
 PS Example 10; Page 82-84; 100pp; English.
 XX

CC The present sequence is the human G-protein G-alpha-16. G-alpha-16
 CC interacts differentially with several receptor types including members of
 CC the opioid and chemokine receptor families. Antisense oligonucleotides
 CC that target the nucleotide sequence encoding the present protein can be
 CC used to modulate the expression of G-alpha-16. They may be used to
 CC inhibit the expression of G-alpha-16 in human cells and tissues and thus
 CC to treat diseases associated with G-alpha-16, such as hyperproliferative
 CC disorders, especially cancer. Infections, inflammation or tumour
 CC formation can be prevented or delayed. The compounds can be
 CC used in research and diagnostics in sandwich and other assays.
 XX
 XX Sequence 374 AA;

Query Match 100.0%; Score 1960; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.6e-189;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MARSLTWRCPCWCLTEDEKAAARVDQENIRILLLEQKKQDRGELKLLLLGPGESGKSTFK 60
 DB 1 marsltwrcpcwcltedekaaarvdqenirillleqkkqdrgeklilllpgesgstfk 60
 QY 61 QMRIIHGAGYSEERKGFPLVYQNIIVSMRAMIEAMERLQIPFSPESKHHASLVMSQD 120
 DB 61 qmriihgagyseerkgfplvyqniivsmramieamerlqipfspeskhhaslvmseqd 120
 QY 121 PYKVTTFEKRYAAAMQWLWRDAGIRACYERRREFHLLDSAVYVLSHLERITEGYVPTAQ 180
 DB 121 pykvttfekryaaamqwlwrdagiracyerrrefhlldsavvylshleriteegyvypta 180
 QY 181 DVLSRMPPTGINEYCFVQKTNLRIVDVGGQKSERKKWHCFENVIALIYLASISEYDQ 240
 DB 181 dvlsrmpptgineycfvqktnlrivdvvggkserkkwhcfenvialiyilasisey 240
 QY 241 CLEENNOENRMKESLALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPSQG 300
 DB 241 cleennqenrmkeslalfgtilelpwfkstsvilfnktdileekiptshlatypsfg 300
 QY 301 PQQDAEAAKRFILDMYTRMYTCVDPGEGSKKGARSRLFSHYTCATDTQNIKRVFKDVR 360
 DB 301 pqqdaeaakrfildmytrmytcvdpgegskkgarsrlfshytcatdtqnirkvfkdv 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 dsvlaryldeinll 374

RESULT 3
ID AAW42434 standard; Protein; 374 AA.
XX AAW42434;
AC AAW42434;
DT 22-JUN-1998 (first entry)
XX Murine promiscuous G-alpha 16 protein.
XX Promiscuous G-alpha 16 protein; mouse; G-protein coupled receptor;
KW ligand; agonist; antagonist; signal transduction.
XX Mus musculus.
XX WO9748820-A1.
XX 24-DEC-1997.
XX 20-JUN-1997; 97WO-US10642.
XX 21-JUN-1996; 96US-0020234.
XX (AURO-) AURORA BIOSCIENCES CORP.
PA Negulescu PA, Offermanns S, Simon M, Zuker C;
XX WPI; 1998-063158/06.
DR N-PSDB; AAV03464.
XX Stable cells containing sequence encoding promiscuous G-alpha
PT protein - useful to identify G-protein coupled receptors or ligands,
PT and agonists or antagonists of signal transduction in cells
XX Disclosure; Page 49-53; 72pp; English.
XX This protein comprises murine G-alpha 16 protein, a promiscuous
CC G-protein whose subunits allow coupling with G-protein coupled
CC receptors (GPCRs) that normally couple with G-proteins of other
CC families. Stable isolated cells are claimed that include a
CC construct comprising an inducible promoter linked to a nucleic
CC acid encoding a promiscuous G-alpha protein. A murine G-alpha 16
CC protein polynucleotide (see AAV03464) or a human G-alpha 15 protein
CC polynucleotide (see also AAV03465) is preferably used. The cells
CC may also comprise a second construct in which a reporter gene is
CC linked to a second promoter that is modulated by a promiscuous
CC G-alpha protein. The cells are used in novel methods for
CC identifying a GPCR for a given ligand, or vice versa, or modulators
CC of signal transduction in a cell and for classifying ligands as
CC agonists or antagonists. Since live cells are used, any identified
CC receptor or ligand can be cloned, and use of fluorescent detection
CC permits characterisation of individual cells.
XX Sequence 374 AA;
SQ
Query Match 99.3%; Score 1947; DB 19; Length 374;
Best Local Similarity 99.2%; Pred. No. 3.3e-186;
Matches 371; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MARSLTWRCPCWCLTEDEKAAARVDQENIRILLEOKKQDRGELKLLLLGPGESGKSTFK 60
DB 1 marsltwrcpcwcltedekaaarvdqenirilleokkqdrgelkllllgpgesgkstfk 60
QY 61 QMRIHAGYSEERKGRFLPYQNYFVSMRAMIFAMERLQIPFSRPSKHHASLVMSQD 120
DB 61 qmrihagyseerkgrflpyqnyfvsramieamerlqipfsrpskhhaslvmssqd 120
QY 121 PKYVTFEKRVAAMQWLWRDAGIACVYRRREFHLLDSAVYLLSHLERITEGVPTAQ 180
DB 121 pykvtftekryaamqwlwrdagiacvyrrefhlldsavylshleriteegvptaq 180

QY 181 DVLRSRMPPTGINEXCFVQKTNLRIVDVGQKSERKKWIHCFENVIALIYLASLSEYDQ 240
DB 181 dvlrsrmpptginexcfvqktnlrivdvggrserkkwi hcfenvialiy laslse ydq 240
QY 241 CLEENNOENRMKESALFCTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFFSFQ 300
DB 241 cleennoenrmkesal fctilelpwfkstsvil flnkt dleekiptshlaty ffsf q 300
QY 301 PRQDAEAAKRFILDMYRMTCVDCGPEGSKGARGSRRLFSHYTCATDTQNIKRVKDV 360
DB 301 prqdaeaakrfildmyrm tcvdcgpegskgarg srrlfs hytc atdtqni krvkdvr 360
QY 361 DSVLARYLDEINLL 374
DB 361 dsvlaryldeinll 374
RESULT 4
ID AAY93970 standard; Protein; 374 AA.
XX AAY93970;
AC AAY93970;
XX 03-OCT-2000 (first entry)
DT Amino acid sequence of a rabbit G-protein alpha 16 polypeptide.
DE Rabbit; G-protein alpha 16; signal transduction; immunogen.
KW Oryctolagus cuniculus.
XX WO200036920-A1.
XX 29-JUN-2000.
XX 21-DEC-1999; 99WO-US30726.
XX 22-DEC-1998; 98US-0218489.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Ames RS, Feild JA, Testa T;
XX WPI; 2000-442490/38.
XX N-PSDB; AAA57483.
XX Novel rabbit G-alpha 16 polypeptides for antibody production and for
PT screening antagonists and agonists of G-coupled protein receptor -
XX Claim 1; Page 26-27; 30pp; English.
XX The present sequence represents a rabbit G-protein alpha 16 polypeptide.
CC The polypeptide is believed to be a member of the GTP-binding regulatory
CC protein family. The polypeptide is a regulator of signal transduction.
CC The polypeptide is useful for treating an individual in need of
CC enhanced activity or expression of G-protein alpha 16. Antagonists are
CC useful for treating an individual in need to inhibit the activity or
CC expression of G-protein alpha 16. The polypeptide is also useful as an
CC immunogen to produce antibodies and for producing immunological response
CC in mammals. The G-protein alpha 16 polypeptides and polynucleotides are
CC useful in screening for antagonists and agonists of G-coupled protein
CC receptors.
XX Sequence 374 AA;
SQ
Query Match 87.3%; Score 1711; DB 21; Length 374;
Best Local Similarity 85.8%; Pred. No. 2.4e-164;
Matches 321; Conservative 29; Mismatches 24; Indels 0; Gaps 0;
QY 1 MARSLTWRCPCWCLTEDEKAAARVDQENIRILLEOKKQDRGELKLLLLGPGESGKSTFK 60
DB 1 marsltwrcpcwcltedekaaarvdqenirilleokkqdrgelkllllgpgesgkstfk 60

Db 1 marslawrcpwcslsedekaaarvdqetrllehhrrqvrgelklilllgtgesgkstfik 60
 -Qy 61 QMRIHAGYSEERKGRPLVQNIIVSMRAMTEAMERLQIPFSPESKHASLVMSOD 120
 Db 61 qmrihagayseerdkgrplvfqnlflsvqailcamdrilqpyrpeskklhaslvmsqd 120
 Qy 121 PYKVTFEKKYAAAMQWLWDAGIRACYERRRHHLLDSAVYLSHLERITEEGYVPTAQ 180
 Db 121 pykvtfefkylavaglsrwdagiracyerrrrhllldsavylshleriteegyvptac 180
 Qy 181 DVLSRMPPTGNEYCFVSQKTNLRIVDVGQSKSERKKWHCFENVIALYLSLSEYDQ 240
 Db 181 dvlsrmpptgineycfsvqtknlrvdvgqskserkwhcfenvialylslsleydq 240
 Qy 241 CLEENQENRMKESLALFGFIPLELPWPKTSVILFLNKTDLIEEKIPTSHLATYFPSPQ 300
 Db 241 cleengqenrmqeslalfgtvlpalpfratsvflfntdiledkvtvrtshlatyfpgr 300
 Qy 301 PKODAAAKRFILDMYTRMYTCVGDGPEGSKKGARSRRFLSHYTCATDTQNIKRVKDV 360
 Db 301 pkodpeaakrfillelytryvagaagdgaskgrsrlfshytcatdtqnikrvkdv 360
 Qy 361 DSVLARYLDEINLL 374
 Db 361 dsvlaryldeinll 374

RESULT 5
 AAY49120
 ID AAY49120 standard; Protein; 374 AA.
 AC AAY49120;
 XX
 DT 07-JAN-2000 (first entry)
 DE G protein alpha 15.
 KW G-protein fusion receptor; Ca; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABAR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 OS Homo sapiens.
 XX
 PN WO9951641-A1.
 XX
 PD 14-OCT-1999.
 XX
 PF 02-APR-1999; 99WO-US07333.
 XX
 PR 03-APR-1998; 98US-0080671.
 XX
 PA (NPSF-) NPS PHARM INC.
 PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 XX
 DR WPI; 1999-610995/52.
 XX
 PT New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease -
 XX
 PS Claim 32; Fig 4; 255pp; English.
 CC
 CC The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a Car
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABAR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (Gp) linked to ICD or the

CC linker. (I), and recombinant chimeric receptors (CR) without the Gp
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used;
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of Car and mGluR
 CC domains allows presentation of GABAR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 CC
 SQ Sequence 374 AA;
 Query Match 87.0%; Score 1705; DB 20; Length 374;
 Best Local Similarity 84.8%; Pred. No. 9.5e-164;
 Matches 317; Conservative 34; Mismatches 23; Indels 0; Gaps 0;
 Qy 1 MARSILTWRCCPWCCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 Db 1 marsiltwccpwcclteeektaaridqeinrilleqkkqereelkllllgpgesgkstfik 60
 Qy 61 QMRIHAGYSEERKGRPLVQNIIVSMRAMTEAMERLQIPFSPESKHASLVMSOD 120
 Db 61 qmrihagayseerdkgrplvfqnlflsvqailcamdrilqpyrpeskklhaslvmtqd 120
 Qy 121 PYKVTFEKKYAAAMQWLWDAGIRACYERRRHHLLDSAVYLSHLERITEEGYVPTAQ 180
 Db 121 pykvtfekypavamqylwrdagiracyerrrrhllldsavylshleriteegyvptaq 180
 Qy 181 DVLSRMPPTGNEYCFVSQKTNLRIVDVGQSKSERKKWHCFENVIALYLSLSEYDQ 240
 Db 181 dvlsrmpptgineycfsvqtknlrvdvgqskserkwhcfenvialylslsleydq 240
 Qy 241 CLEENQENRMKESLALFGFIPLELPWPKTSVILFLNKTDLIEEKIPTSHLATYFPSPQ 300
 Db 241 cleengqenrmeslalfstlelpwfktsvllfntdiledkhtshlatyfpfsg 300
 Qy 301 PKODAAAKRFILDMYTRMYTCVGDGPEGSKKGARSRRFLSHYTCATDTQNIKRVKDV 360
 Db 301 prdaaakrfildmyarvyascaeppdqggrkgsrarrffahfctatdtqsvrvfkdv 360
 Qy 361 DSVLARYLDEINLL 374
 Db 361 dsvlaryldeinll 374

RESULT 6
 AAW42435
 ID AAW42435 standard; Protein; 374 AA.
 AC AAW42435;
 XX
 DT 22-JUN-1998 (first entry)
 DE Human promiscuous G-alpha 15 protein.
 XX
 KW Promiscuous G-alpha 15 protein; human; G-protein coupled receptor;
 KW ligand; agonist; antagonist; signal transduction.
 XX
 OS Homo sapiens.
 XX
 PN WO9748820-A1.
 XX
 PD 24-DEC-1997.
 XX
 PF 20-JUN-1997; 97WO-US10642.
 XX

QY 133 AAMQWLWDAGIRACRYERREFFHLLDSAVYVLSHLERITEEGYVPTAODVLSRMPPTGI 192
 Db 1047 daikslwvdpqicqcyrrreyqlsdsdkyindldrvadpalyptqgdvrvrvptgti 1106
 QY 193 NEYCFVSQNTNLRIYDVGQSKSRKWHCFENVIALIYASLSEYDOCLEENNOENRMK 252
 Db 1107 leypfdlqsvifrmvvggrserrkwhcfenvtslmfvalseydqvivesdnenme 1166
 QY 253 ESALFGTILELPWFKSTSVILEFNKTDILEEKIPTSHLATYPPSFGPKQDAEAAKRFI 312
 Db 1167 eskalftrtiitypwwfnqssvillfnkdlleekimyslhvdyfpeydgprdaqarefi 1226
 QY 313 LDMYTRMTGCVDPGEGSKGARSRLFSHYTCATDTONIRKVKFDVRSVLARYLDEIN 372
 Db 1227 lkmfvdl-----npsdk-----liyshftcatdenirfvfaavkdtlqnlkdcg 1274
 QY 373 L 373
 Db 1275 L 1275
 RESULT 8
 AAY49129
 ID AAY49129 standard; Protein; 1394 AA.
 XX
 AC AAY49129;
 XX
 DT 07-JAN-2000 (first entry)
 DE pmGluR2/Car*Galpha15 fusion construct protein sequence.
 XX
 KW G-protein fusion receptor; Car; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 XX
 OS Homo sapiens.
 OS WO9951641-A1.
 PN 14-OCT-1999.
 XX
 PF 02-APR-1999; 99WO-US07333.
 XX
 PR 03-APR-1998; 98US-0080671.
 XX
 PA (NPSF-) NPS PHARM INC.
 XX
 PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 XX
 DR WPI; 1999-610995/52.
 DR N-PSDB; AA231060.
 XX
 PT New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease -
 XX
 PS Example 1; Fig 12; 255pp; English.
 XX
 CC The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a Car
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
 CC linker. (I), and recombinant chimeric receptors (CR) without the GP
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or

CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (1) is used:
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of Car and mGluR
 CC domains allows presentation of GABABR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 XX
 SQ Sequence 1394 AA;
 Query Match 52.9%; Score 1337; DB 20; Length 1394;
 Best Local Similarity 55.4%; Pred. No. 2.2e-95;
 Matches 200; Conservative 61; Mismatches 88; Indels 12; Gaps 2;
 QY 13 CLTEDEKAAARVDOEINRILLEQKKODRGELKLLILGPGESGKSTFIKOWRIIHGAGYSE 72
 Db 1045 clseakearrindelqrldrdarfelkllilgtgesgkstfkgmrlilngsgysd 1104
 QY 73 EERGFRLVYQNIYVSMRAMTEAMERLOIP:SRPESKHASLVMSQDPKYKVTTFEKRYA 132
 Db 1105 edkrgftklvyqnlftamqamramdtklipykvehnkahqqlrvrevdveksafenpyv 1164
 QY 133 AAMQWLWDAGIRACRYERREFFHLLDSAVYVLSHLERITEEGYVPTAODVLSRMPPTGI 192
 Db 1165 daikslwvdpqicqcyrrreyqlsdsdkyindldrvadpalyptqgdvrvrvptgti 1224
 QY 193 NEYCFVSQNTNLRIYDVGQSKSRKWHCFENVIALIYASLSEYDOCLEENNOENRMK 252
 Db 1225 leypfdlqsvifrmvvggrserrkwhcfenvtslmfvalseydqvivesdnenme 1284
 QY 253 ESALFGTILELPWFKSTSVILEFNKTDILEEKIPTSHLATYPPSFGPKQDAEAAKRFI 312
 Db 1285 eskalftrtiitypwwfnqssvillfnkdlleekimyslhvdyfpeydgprdaqarefi 1344
 QY 313 LDMYTRMTGCVDPGEGSKGARSRLFSHYTCATDTONIRKVKFDVRSVLARYLDEIN 372
 Db 1345 lkmfvdl-----npsdk-----liyshftcatdenirfvfaavkdtlqnlkdcg 1392
 QY 373 L 373
 Db 1393 L 1393
 RESULT 9
 AAY49134
 ID AAY49134 standard; Protein; 1397 AA.
 XX
 AC AAY49134;
 XX
 DT 07-JAN-2000 (first entry)
 DE pmGluR2/Car*Galpha15+3Ala linker fusion construct protein sequence.
 XX
 KW G-protein fusion receptor; Car; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 XX
 OS Homo sapiens.
 OS WO9951641-A1.
 PN 14-OCT-1999.
 PD 02-APR-1999;
 PF 02-APR-1999; 99WO-US07333.
 XX
 PR 03-APR-1998; 98US-0080671.

RESULT 10

AAV49131 standard. protein: 1418 AA.

ID AAY49131
XX
AC AAY49131:

07-JAN-2000 (first entry)

XX - <http://www.ncbi.nlm.nih.gov/Genbank/Genbank.html>

XX G-protein fusion receptor; CAR; calcium receptor; GluR; head injury;
DE metabotropic glutamate receptor; GABA_BR; chimeric receptor; stroke;
XX gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW cognitive disorder.

XX
XX
XX

OS Homo sapiens.
XX

PN WO9951641-A1
XX

PD 14-OCT-1999.
XX

02-APR-1999; 99WO-US07333.

AA
PR 03-APR-1998; 98US-00XX
XX
PA (NPSP-) NPS PHARM INC.XX
PI Stormann TM, HammerlPI Simin RT;
yy

DR WPI: 1999-610995/52.
DR N-PSDB; AA231062.
XX
XX New G-protein fusion receptors and chimeras containing domains from
PT different receptors, used to screen for modulators, potentially useful
PT for the treatment of Alzheimer's disease.

PT e.g. for treating or preventing str

Example 1; Fig 12; 255pp; English.

The invention relates to G-protein coupled fusion receptors (I) comprising: (1) in the N to C direction, extracellular (EC), transmembrane (TMD) and intracellular (ICD) domains, each chosen independently from a CaR (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR (gamma-aminobutyric acid receptor); (2) an optional linker attached to the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the linker. (1), and recombinant chimeric receptors (CR) without the GP component, are used to assess function of the various domains and to identify compounds (e.g. allosteric modulators or antagonists) that act on these domains. The modulators are potentially useful for treating or preventing diseases associated with the receptors, e.g. stroke, head or spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia, Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive disorders and depression. Nucleic acid (II) that encodes (I) is used: (1) for recombinant production of corresponding proteins; and (2) to produce cells used in screening for modulators. Use of CaR and mGluR domains allows presentation of GABABR domains, to a binding agent, in a form more like the natural domain structure compared with use of incomplete receptors, lacking one or more domains. By shuffling different domains, agents can be identified that affect particular domains of a receptor.

XX	Sequence	1418 AA;
SO		

Query Match	52.9%	Score 1037	DB 20	Length 1418
-------------	-------	------------	-------	-------------

[illegible]

(NPSP-) NPS PHARM INC.
Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
Simin RT;
WPI; 1999-610995/52.
N-PSDB; AA231065.
New G-protein fusion receptors and chimeras containing domains from
different receptors, used to screen for modulators, potentially useful
e.g. for treating or preventing stroke or Alzheimer's disease -
Example 1: Fig 12; 255pp; English.

AX	Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
IX	

Simin RT;

WPI; 1999-610995/5

New G-protein fusion receptors and chimeras containing domains from different receptors, used to screen for modulators, potentially useful e.g. for treating or preventing stroke or Alzheimer's disease - Example 1; Fig 12; 255pp; English.

the invention relates to G-protein

the involvement of the extracellular (ECD), transmembrane (TMD) and intracellular (ICD) domains, each chosen independently from a CaR (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR (gamma-aminobutyric acid receptor). (2) an optional linker attached to the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the linker. (1), and recombinant chimeric receptors (CR) without the GP component, are used to assess function of the various domains and to identify compounds (e.g. allosteric modulators or antagonists) that act on these domains. The modulators are potentially useful for treating or preventing diseases associated with the receptors, e.g. stroke, head or spinal cord injury, epilepsy, ischemia, hypoglycaemia, anoxia, Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive disorders and depression. Nucleic acid (II) that encodes (1) is used: (1) for recombinant production of corresponding proteins; and (2) to produce cells used in screening for modulators. Use of CaR and mGluR domains allows presentation of GABABR domains, to a binding agent, in a form more like the natural domain structure compared with use of incomplete receptors, lacking one or more domains. By shuffling different domains, agents can be identified that affect particular domains of a receptor.

XX	
SO	Sequence 1397 AA;

Query Match	52.9%	Score 1037;	DB 20;	Length 1397;
Best Local Similarity	55.4%	Pred. No. 2.2e-95;		
Matches 200;	Conservative	61;	Mismatches 88;	Indels 12;
Gaps 2;				

QY 13 CLTEDEKAAARVDQEINRILLEQKKQDORGELKLLLPGESCKSTFIKQMRIIHGACYSE 72

```
Db 1048 clseeaxearrndeierqlrrdkdarrelkl|||gtgesgkstfikqmriihgsysd 1107
```

73 EERKGERPLVYQNTVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTEKRYA 132

1108 okkrqftv lvvgn ift amgami ramdt lk pkyvehnkahaq lvrevdveksafenpyv 1167

DB 1106 eanrigckrivyqurazcamfemamrnfjff
-
122 rancocwrrdnactpacvcepppeehiidsavvyishleriteegvvptoadvlrsnpttgi 192

QY I33 AAMQWLRDASLRACLEKREKHEEDDAVYXZBONHAKRAT 1227
{: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1168 dalkslwndpbigecyarreyqisdskyyinurduvaupoyipdqquwvllvwpccy- 250

QY 193 NEYCFSVQKTNLRIVDVGQKSERKKWIIHCFENFIALIYLASTSEYDQCLENNQENRKN 252

Db 1228 ieypdfqlsqvifrmvmdvggqrserkwhicfenvtsimflvalseydqvivesdnrme 128/

253 ESALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSGQPKQDAEAAKRFI 312

```
Db 1288 eska!frtiitypwfqgnssv!flnkkd!leekimyslvdypdygpprdqaarefi 1347
```

313 LDMYTRMYTGCVDGPECKKGGARSRRLLFSHYTCATDTQIRKVKFQDVRDSVLARLDEIN 372

1348 1 km fvd

[illegible]

QY 73 EERKGRPLVYQNIYFVMSRAMIEMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 132
 Db 1129 edkrgftklvygnifcamqamiramdtklipykayehnkahaglvrevdvkvsafepv 1188
 QY 133 AAMQWLDRDAGIRACYERREFFHLLDSAVYLSHLERITEEGYVPTAQDVLKSRMPTGI 192
 Db 1189 daikslwvpgiqecydrreyqsdstkyindlradvapylptqgdlrvvptgti 1248
 QY 193 NEYCFVQKTNLRIVDVGGOKSERKKWHCFENVTALIVLASLSEYDQCLEENQENRMK 252
 Db 1249 leypfdlqsvifrmvvgqqrserkwhcfenvtalsimflvalseydqvivesdnenrme 1308
 QY 253 ESALFGRITLPEWFKSTSVILFLNKTDILEKIPTSHLATYFPSPQGPQDAEAKRFI 312
 Db 1309 eskalftrtiitypwnsqsvllflnkdlleeklmshlvdypdydgqdaqaarefi 1368
 QY 313 LDMYTRMYTCVGDGPEGSKGARSRLFSHYTCATDTONIRKVKDVRSDVSLARYLDEIN 372
 Db 1369 lkmfvdl-----npsdk-----liyshtcatdenirfvfaavkdtllqlnkdcg 1416
 QY 373 L 373
 Db 1417 L 1417

RESULT 11
 AAY49133
 ID AAY49133 standard; Protein; 1323 AA.
 AC AAY49133;
 DT 07-JAN-2000 (first entry)
 DE GABA-BR1a*Gqo5 fusion construct protein sequence.
 KW G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABAB; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 XX Homo sapiens.
 OS
 XX WO9951641-A1.
 PN 14-OCT-1999.
 PD
 PF 02-APR-1999; 99WO-US07333.
 PR 03-APR-1998; 98US-0080671.
 XX (NPSp-) NPS PHARM INC.
 XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 XX WPI: 1999-610995/52.
 DR N-P5DB; AAZ31064.
 XX
 XX New G-protein fusion receptors and chimeras containing domains from
 XX different receptors, used to screen for modulators, potentially useful
 XX e.g. for treating or preventing stroke or Alzheimer's disease -
 XX Disclosure; Fig 14; 255pp; English.
 XX
 XX The invention relates to G-protein fusion receptors (1) comprising:
 XX (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 XX intracellular (ICD) domains, each chosen independently from a CaR
 XX (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
 XX (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 XX the C-terminus of ICD; and (3) a G-protein (Gp) linked to ICD or the

CC linker. (1), and recombinant chimeric receptors (CR) without the GP
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Acetic acid (II) that encodes (I) is used:
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening of modulators. Use of CaR and mGluR
 CC domains allows presentation of GABAB domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 XX Sequence 1323 AA;
 QY
 Query Match 52.8%; Score 1035.5; DB 20; Length 1323;
 Best Local Similarity 55.4%; Pred No. 2.9e-95;
 Matches 200; Conservative 62; Mismatches 82; Indels 17; Gaps 3;
 QY 13 CLTEDEKAAARVDQEIINRILLFQKODRGELKLLLLGPGESGKSTFIKQMRITIHGAGYSE 72
 Db 974 clseeakearrindeierqlrrdkdrareklllllgsegskcfikqmrllhngsydsd 1033
 QY 73 EERKGRPLVYQNIYFVMSRAMIEMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 132
 Db 1034 edkrgftklvygnifcamqamiramdtklipykayehnkahaglvrevdvkvsafepv 1093
 QY 133 AAMQWLDRDAGIRACYERREFFHLLDSAVYLSHLERITEEGYVPTAQDVLKSRMPTGI 192
 Db 1094 daikslwvpgiqecydrreyqsdstkyindlradvapylptqgdlrvvptgti 1153
 QY 193 NEYCFVQKTNLRIVDVGGOKSERKKWHCFENVTALIVLASLSEYDQCLEENQENRMK 252
 Db 1154 leypfdlqsvifrmvvgqqrserkwhcfenvtalsimflvalseydqvivesdnenrme 1213
 QY 253 ESALFGRITLPEWFKSTSVILFLNKTDILEKIPTSHLATYFPSPQGPQDAEAKRFI 312
 Db 1214 eskalftrtiitypwnsqsvllflnkdlleeklmshlvdypdydgqdaqaarefi 1273
 QY 313 LDMYTRMYTCVGDGPEGSKGARSRLFSHYTCATDTONIRKVKDVRSDVSLARYLDEIN 372
 Db 1274 lkmfvdl-----npsdk-----liyshtcatdenirfvfaavkdtll-----qin 1316
 QY 373 L 373
 Db 1317 L 1317

RESULT 12
 AAY49132
 ID AAY49132 standard; Protein; 1303 AA.
 AC AAY49132;
 DT 07-JAN-2000 (first entry)
 DE GABA-BR2*Gqo5 fusion construct protein sequence.
 KW G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABAB; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 XX Homo sapiens.
 OS
 XX WO9951641-A1.
 PN
 XX


```

Db 10 clsdvkeskrinaeiekqlrrdrdrrelkllllgtgesgskstfrikqmriihgagye 69
Qy 73 EERKGRPLVYQNIYFVSMRAMIEMERLOIPFSPRSKHHASLVMSODPKYKVTTEKRYA 132
Db 70 edkrgtklvyqnficamqamirametkllilyeqnkanallirevdekvttfehqv 129
Qy 133 AAMQWLWRDAGIRACYERREPHLLDSAVYVLSHLERITEEGYVPTAODVLRSMPTTGI 192
Db 130 saiktledpglqecydrreyqlsdsakylldvdratiglyptqddvrvrvptgtl 189
Qy 193 NEYCFVQKTNLRIVDVGGOKSERKWHCFENVIALIYLASEYDQCLEENNQENRMK 252
Db 190 ieypfalenlfrmdvvggrrserfkwihcfenvtsimflvalseydqvlivesdnenme 249
Qy 253 ESALFETILELPWFKSTVILFLNKTDILEKIPSHLATYPPSTOGPKODAAKRFI 312
Db 250 eskalfititlypwnssvllfinkdillekilyshlvdyfpefddgqrdqaaarefi 309
Qy 313 LDMYTRMYTCVGDGPGSKGARSRLFSHYTCATDTONIRKVKDVRDVLARYLDEIN 372
Db 310 pkmfvd-----iyshtcatdenirfvaavkdtllqlnkeyn 357
Qy 373 LL 374
Db 358 LV 359

RESULT 14
AAV29789
ID AAV29789 standard; Protein; 359 AA.
XX
AC AAV29789;
XX
DT 15-NOV-1999 (first entry)
XX
DE Human G-alpha-11 protein sequence.
XX
KW Human; G-alpha-11; antisense oligonucleotide; inhibition; expression;
KW phosphorothioate.
XX
OS Homo sapiens.
XX
PN US5951455-A.
XX
PD 14-SEP-1999.
XX
PF 04-DEC-1998; 98US-0205922.
XX
PR 04-DEC-1998; 98US-0205922.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Cowsert LM;
XX
DR WPI; 1999-539140/45.
XX
DR N-PSDB; AAZ19461.
XX
PT Inhibitory antisense compounds useful for the treatment of diseases
PT associated with G-alpha-11
XX
PS Example 13; Column 45-48; 38pp; English.
XX
CC The present invention describes inhibitory antisense compounds of 8-30
CC nucleotides, targeted to a nucleic acid molecule encoding human
CC G-alpha-11. The present sequence represents human G-alpha-11. AAZ19468
CC to AAZ19547 represent human G-alpha-11 phosphorothioate antisense
CC oligonucleotides given in the present invention. The oligonucleotides
CC may be useful for the treatment of diseases associated with G-alpha-11.
XX
SQ Sequence 359 AA;

```

```

Query Match 52.7%; Score 1333; DB 20; Length 359;
Best Local Similarity 55.5%; Pred. No. 7.2e-96;
Matches 201; Conservative 65; Mismatches 84; Indels 12; Caps 2;

Qy 13 CLTEDEKAAARVOENRILLEOKKQDRGBELKLLLGPGESGKSTFIKQMRIIHGAGYSE 72
Db 10 clsdvkeskrinaeiekqlrrdrdrrelkllllgtgesgskstfrikqmriihgagye 69
Qy 73 EERKGRPLVYQNIYFVSMRAMIEMERLOIPFSPRSKHHASLVMSODPKYKVTTEKRYA 132
Db 70 edkrgtklvyqnficamqamirametkllilyeqnkanallirevdekvttfehqv 129
Qy 133 AAMQWLWRDAGIRACYERREPHLLDSAVYVLSHLERITEEGYVPTAODVLRSMPTTGI 192
Db 130 saiktledpglqecydrreyqlsdsakylldvdratiglyptqddvrvrvptgtl 189
Qy 193 NEYCFVQKTNLRIVDVGGOKSERKWHCFENVIALIYLASEYDQCLEENNQENRMK 252
Db 190 ieypfalenlfrmdvvggrrserfkwihcfenvtsimflvalseydqvlivesdnenme 249
Qy 253 ESALFETILELPWFKSTVILFLNKTDILEKIPSHLATYPPSTOGPKODAAKRFI 312
Db 250 eskalfititlypwnssvllfinkdillekilyshlvdyfpefddgqrdqaaarefi 309
Qy 313 LDMYTRMYTCVGDGPGSKGARSRLFSHYTCATDTONIRKVKDVRDVLARYLDEIN 372
Db 310 pkmfvd-----iyshtcatdenirfvaavkdtllqlnkeyn 357
Qy 373 LL 374
Db 358 LV 359

RESULT 15
AAV49125
ID AAV49125 standard; Protein; 359 AA.
XX
AC AAV49125;
XX
DT 07-JAN-2000 (first entry)
XX
DE Chimeric Gq15 protein sequence.
XX
KW G-protein fusion receptor; CaR; calcum receptor; GluR; head injury;
KW metabotropic glutamate receptor; GABAAR; chimeric receptor; stroke;
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW cognitive disorder.
XX
OS Homo sapiens.
XX
PA WO9951641-A1.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07333.
XX
PR 03-APR-1998; 98US-0080671.
XX
PA (NPS- ) NPS PHARM INC.
XX
PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI Simlin RI;
XX
DR WPI; 1999-610995/52.
DR N-PSDB; AAZ31056.
XX
PT New G-protein fusion receptors and chimeras containing domains from
PT different receptors, used to screen for modulators, potentially useful
PT e.g. for treating or preventing stroke or Alzheimer's disease
XX
PS Disclosure; Page 196-197; 255pp; English.

```

Sequence 359 AA;

search completed: September 6, 2001, 10:54:33
Job time: 53 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 6, 2001, 10:53:40 ; Search time 13.15 Seconds
(without alignments)
974.262 Million cell updates/sec

Title: US-08-878-801-2
Perfect score: 1960
Sequence: 1 MARSLTRCCPWCLTEDEKA.....VFKDVRDSVLARYLDEINLL 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1960	100.0	374	1 GB15_HUMAN	P30679 homo sapien
2	1722	87.9	374	1 GB15_HUMAN	O88302 rattus norv
3	1705	87.0	374	1 GB15_MOUSE	P30678 mus musculu
4	1052	53.7	353	1 GBQ_XENLA	P38410 xenopus lae
5	1047	53.4	353	1 GBQ_CANFA	Q28294 canis fami
6	1046	53.4	353	1 GBQ_RAT	P82471 rattus norv
7	1046	53.4	359	1 GB11_BOVIN	P38409 bos taurus
8	1044	53.3	353	1 GBQ_MOUSE	P21279 mus musculu
9	1042	53.2	353	1 GBQ_HUMAN	P50148 homo sapien
10	1040	53.1	359	1 GB11_MOUSE	P29992 homo sapien
11	1038	53.0	359	1 GB11_MOUSE	P21278 mus musculu
12	1036	52.9	353	1 GBQ_HOMAM	P91950 homarus ame
13	1035	52.8	359	1 GB11_MELGA	P45645 melagris g
14	1030	52.6	359	1 GB11_RAT	Q9J102 rattus norv
15	1026.5	52.4	359	1 GBQ_XENLA	P43444 xenopus lae
16	1026	52.3	353	1 GBQ1_DROME	P23625 drosophila
17	1026	52.3	353	1 GBQ3_DROME	P54400 drosophila
18	1017.5	51.9	354	1 GBQ_LOLFO	P38412 loligo forb
19	1013	51.7	354	1 GB14_XENLA	O73819 xenopus lae
20	1013	51.7	355	1 GB14_BOVIN	P38408 bos taurus
21	1013	51.7	355	1 GB14_MOUSE	P30677 mus musculu
22	1005	51.3	355	1 GB14_HUMAN	O95837 homo sapien
23	995	50.8	353	1 GBQ_LYMST	P38411 lymnaea sta
24	989	50.5	353	1 GBQ_PATYE	O15975 patinopecte
25	787.5	40.2	352	1 GBAL_EMENI	Q00743 emericella
26	766.5	39.1	353	1 GBAL_NEUCR	Q05425 neurospora
27	764	39.0	352	1 GBAL_COCHE	O74227 cochliobolu
28	762	38.9	354	1 GB02_DROME	P16377 drosophila
29	760.5	38.8	354	1 GB01_DROME	P16378 drosophila
30	760	38.8	352	1 GBAL_COPCO	P30675 coprinus co
31	759.5	38.8	352	1 GBAL_CRYPA	Q00580 cryphonectr
32	759.5	38.8	353	1 GBAL_USYMA	P87032 ustilago ma
33	757.5	38.6	352	1 GBAL_COLTR	O42784 colletotric

RESULT 1

ID	GB15_HUMAN	STANDARD;	PRT;	374 AA.
AC	P30679; O75247;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-15 SUBUNIT (ALPHA-16).			
GN	GNA15 OR GNA16.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91288509; PubMed=1905813;			
RA	Amatruda T.T. III, Steele D.A., Slepak V.Z., Simon M.I.;			
RT	"G alpha 16, a G protein alpha subunit specifically expressed in hemopoietic cells.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:5587-5591(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,			
RA	Burkhardt-Schultz K., Gordon L., Ramirez M., Stilwagen S.,			
RA	Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,			
RA	Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,			
RA	Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,			
RA	Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,			
RA	Krommiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,			
RA	Kobayashi A., Olsen A.S., Carrano A.V.;			
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
CC	FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS.			
CC	SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA). THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.			
CC	TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN HEMATOPOIETIC CELLS.			
CC	SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G10).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
DR	EMBL; M63904; AAA35860.1;			
DR	EMBL; AC005264; AAC25612.1;			
DR	EMBL; AC005262; AAC25616.1;			
DR	PIR; A41096; A41096.			
DR	HSP; P04896; IAZT.			
DR	MIM; 139314;			
DR	InterPro; IPR000654;			
DR	InterPro; IPR001019;			
DR	Pfam; PF00503; G-alpha; 1.			

ALIGNMENTS

DR PRINTS; PR00318; GPROTEINA.
 DR PRINTS; PR00442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 FT NP_BIND 49 56 GTP (BY SIMILARITY).
 FT NP_BIND 208 212 GTP (BY SIMILARITY).
 FT NP_BIND 277 280 GTP (BY SIMILARITY).
 FT MOD_RES 186 186 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT CONFLICT 147 147 C -> Y (IN REF. 2).
 SQ SEQUENCE 374 AA; 43508 MW; 8127AC16FC212507 CRC64;

Query Match 100.0%; Score 1960; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred No. 2.9e-150;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLTWRCPCWCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 DB 1 MARSLTWRCPCWCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 QY 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSRPEKHHASLVMSQD 120
 DB 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSRPEKHHASLVMSQD 120
 QY 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVVYLSHLERITEGYVPTAQ 180
 DB 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVVYLSHLERITEGYVPTAQ 180
 QY 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVVYLSHLERITEGYVPTAQ 180
 DB 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVVYLSHLERITEGYVPTAQ 180
 QY 181 DVLSRMPTTGINEXYCFVQKTNLRIVDVGGGKSERKKWIHCFFENVIATLYLASLSEYDQ 240
 DB 181 DVLSRMPTTGINEXYCFVQKTNLRIVDVGGGKSERKKWIHCFFENVIATLYLASLSEYDQ 240
 QY 241 CLEENNOENRMKESLALFGLTLELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 DB 241 CLEENNOENRMKESLALFGLTLELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 QY 301 PKQDAEAAKRFILDMYTRMYTGVCDGPEGSKGARSRRFLSHYTCATDTQNIKKVKFDVR 360
 DB 301 PKQDAEAAKRFILDMYTRMYTGVCDGPEGSKGARSRRFLSHYTCATDTQNIKKVKFDVR 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 DSVLARYLDEINLL 374

RESULT 2
 GB15_RAT
 ID GB15_RAT STANDARD; PRT: 374 AA.
 AC GB15_RAT
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-15 SUBUNIT.
 GN GNA15.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98351893; PubMed=9685675;
 RA Kusakabe Y., Yamaguchi E., Tanemura K., Kaneyama K., Chiba N.,
 RA Arai S., Emori Y., Abe K.,
 RT "Identification of two alpha-subunit species of GTP-binding proteins,
 RT G alpha 15 and G alpha 4, expressed in rat taste buds.";
 RL Biochim. Biophys. Acta 1403:265-272(1998).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB015308; BAA28927.1; -
 DR HSSP; P04896; IAZT.
 DR InterPro; IPR000634; -
 DR InterPro; IPR001019; -
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR PRINTS; PR00442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 FT NP_BIND 49 56 GTP (BY SIMILARITY).
 FT NP_BIND 208 212 GTP (BY SIMILARITY).
 FT NP_BIND 277 280 GTP (BY SIMILARITY).
 FT MOD_RES 186 186 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT CONFLICT 147 147 C -> Y (IN REF. 2).
 SQ SEQUENCE 374 AA; 43331 MW; E5E65D32D41EC0EB CRC64;

Query Match 87.9%; Score 1122; DB 1; Length 374;
 Best Local Similarity 86.1%; Pred No. 3.8e-131;
 Matches 322; Conservative 30; Mismatches 22; Indels 0; Gaps 0;

QY 1 MARSLTWRCPCWCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 DB 1 MARSLTWRCPCWCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 QY 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSRPEKHHASLVMSQD 120
 DB 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSRPEKHHASLVMSQD 120
 QY 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVVYLSHLERITEGYVPTAQ 180
 DB 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVVYLSHLERITEGYVPTAQ 180
 QY 181 DVLSRMPTTGINEXYCFVQKTNLRIVDVGGGKSERKKWIHCFFENVIATLYLASLSEYDQ 240
 DB 181 DVLSRMPTTGINEXYCFVQKTNLRIVDVGGGKSERKKWIHCFFENVIATLYLASLSEYDQ 240
 QY 241 CLEENNOENRMKESLALFGLTLELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 DB 241 CLEENNOENRMKESLALFGLTLELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 QY 301 PKQDAEAAKRFILDMYTRMYTGVCDGPEGSKGARSRRFLSHYTCATDTQNIKKVKFDVR 360
 DB 301 PKQDAEAAKRFILDMYTRMYTGVCDGPEGSKGARSRRFLSHYTCATDTQNIKKVKFDVR 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 DSVLARYLDEINLL 374

RESULT 3
 GB15_MOUSE
 ID GB15_MOUSE STANDARD; PRT: 374 AA.
 AC P30678;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-15 SUBUNIT.
 GN GNA15 OR GNA-15.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92052208; PubMed=1946421;

RA Wilkie T.M., Scherle P.A., Strathmann M.P., Slepak V.Z., Simon M.I.;
 RT "Characterization of G-protein alpha subunits in the Gq class:
 RT expression in murine tissues and in stromal and hematopoietic cell
 RT lines";
 RT Proc. Natl. Acad. Sci. U.S.A. 88:10049-10053(1991).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=129/SV;
 RC MEDLINE=96435432; PubMed=8838318;
 RX Davignon I., Barnard M., Gavrilova O., Sweet K.K., Wilkie T.M.;
 RA "Gene structure of murine Gna11 and Gna15: tandemly duplicated Gq
 RT class G protein alpha subunit genes.";
 RT Genomics 31:359-366(1996).
 RL
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC
 CC -1- TISSUE SPECIFICITY: HEMATOPOIETIC CELLS.
 CC
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M80632; AAA37713.1; -;
 DR EMBL; U37419; AAB36840.1; -;
 DR EMBL; U37414; AAB36840.1; JOINED.
 DR EMBL; U37415; AAB36840.1; JOINED.
 DR EMBL; U37416; AAB36840.1; JOINED.
 DR EMBL; U37417; AAB36840.1; JOINED.
 DR EMBL; U37418; AAB36840.1; JOINED.
 DR PIR; B41534; B41534.
 DR HSSP; P04896; IAZT.
 DR MGD; MGI:95770; Gna15.
 DR InterPro; IPR000654; -;
 DR InterPro; IPR001019; -;
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR PRINTS; PR00442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 FT NP_BIND 49 56 GTP (BY SIMILARITY).
 FT NP_BIND 208 212 GTP (BY SIMILARITY).
 FT NP_BIND 277 280 GTP (BY SIMILARITY).
 FT MOD_RES 186 186 ADP-RIBOSYL(1) (BY ACTION OF CTX)
 FT (BY SIMILARITY).
 FT
 FT SEQUENCE 374 AA; 43535 MW; 44F15DC52C8C233E CRC64;
 SQ
 Query Match 87.08; Score 1705; DB 1; Length 374;
 Best Local Similarity 84.88; Pred. No. 8.7e-130;
 Matches 317; Conservative 34; Mismatches 23; Indels 0; Gaps 0;
 QY 1 MARSLTWCCPWLCTEDEFKAAARVDQENRILLEGKQDQKDELKLLGPGESGKSTFIK 60
 DB 1 MARSLTWCCPWLCTEDEFKAAARVDQENRILLEGKQDQKDELKLLGPGESGKSTFIK 60
 - QY 61 QMRIIHGAGYSEERKGRPLVYQNIYFVSMRAMIDAMERLQIPFSPRESKHHASLVMSQD 120
 DB 61 QMRIIHGAGYSEERKGRPLVYQNIYFVSMRAMIDAMERLQIPFSPRESKHHASLVMSQD 120
 QY 121 PKVVTTFEKRVAAMQWLWRDAGIRACVRRREFHLLDSAVYYLSHLERITEEGVPTPAQ 180
 DB 121 PKVVTTFEKRVAAMQWLWRDAGIRACVRRREFHLLDSAVYYLSHLERITEEGVPTPAQ 180
 QY 121 PKVSTFEKPYAVAMQYLWRDAGIRACVRRREFHLLDSAVYYLSHLERISEDSVPTPAQ 180
 DB 121 PKVSTFEKPYAVAMQYLWRDAGIRACVRRREFHLLDSAVYYLSHLERISEDSVPTPAQ 180
 QY 181 DVLRSRMTPTTGNEVCFVSQKTNLRIVDVGOKSRKKWIKCFENVIALIYLSLSEYDQ 240
 DB 181 DVLRSRMTPTTGNEVCFVSQKTNLRIVDVGOKSRKKWIKCFENVIALIYLSLSEYDQ 240

QY 241 CLENNQENRMKESLALFGLTLELPWFKSTSVILFLNKTIDILEKIPTSHLATYFPSPQ 300
 DB 241 CLENNQENRMKESLALFGLTLELPWFKSTSVILFLNKTIDILEKIPTSHLATYFPSPQ 300
 QY 301 PKQDAEAAKRFILDMYTRMTGCVDPGEGSKKGARSRLSHYTCATDQTNIRKVEKDV 360
 DB 301 PRDAEAAKSFILDMYARVYASCAEPQDGGKGRARRRFFAHFTCATDQTSVRSEKDV 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 DSVLARYLDEINLL 374
 RESULT 4
 GEO_XENLA STANDARD; PRT; 353 AA.
 AC P38410; 1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(Q), ALPHA SUBUNIT.
 GN GNAQ.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Ladner K.J., Smith L.D.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=94298961; PubMed=8026589;
 RA Shapira H., Way J., Lipinsky D., Oron Y., Battay J.F.;
 RT "Neuromedin B receptor, expressed in Xenopus laevis oocytes,
 RT selectively couples to G alpha q and not G alpha 11.";
 RL FEBS Lett. 348:89-92(1994).
 RN [3]
 RN ERRATUM.
 RA Shapira H., Way J., Lipinsky D., Oron Y., Battay J.F.;
 RL FEBS Lett. 349:318-318(1994).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L05540; AAA49730.1; ALT_INIT.
 DR EMBL; U10502; AAA52189.1; ALT_INIT.
 DR HSSP; P04896; IAZT.
 DR InterPro; IPR000654; -;
 DR InterPro; IPR001019; -;
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR PRINTS; PR00442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
 FT NP_BIND 3 3 PALMITATE (BY SIMILARITY).
 FT LIPID 4 4 PALMITATE (BY SIMILARITY).
 FT NP_BIND 40 47 GTP (BY SIMILARITY).
 FT NP_BIND 199 203 GTP (BY SIMILARITY).
 FT NP_BIND 268 271 GTP (BY SIMILARITY).

FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT CONFLICT 10 10
 FT CONFLICT 79 79 E -> K (IN REF. 2).
 FT CONFLICT 158 158 S -> T (IN REF. 2).
 FT CONFLICT 164 164 L -> V (IN REF. 2).
 FT CONFLICT 164 164 H -> Q (IN REF. 2).
 SQ SEQUENCE 353 AA; 41541 MW; 2310142412184553 CRC64;

Query Match 53.7%; Score 1052; DB 1; Length 353;
 Best Local Similarity 56.4%; Pred. No. 2.3e-77;
 Matches 204; Conservative 60; Mismatches 86; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDQENRILLOKODRGELKLLGLGSGSGKSTFTKQRIHAGYSE 72
 DB 4 CLSEAEAEARRINDEIERLRDRKROARRELKLLGLTGSGSGKSTFTKQRIHAGYSE 63
 QY 73 EERKGRPLVYQNIYFVSMRAMIEMERLQIPSPRPSKHHASLVMSODPKYVTFEKRYA 132
 DB 64 EDKRGFTKLVQNIYFAMQAMIRAMDTLKIP:KYEHKHAQHLVREVDVEKVSFENPVY 123
 QY 133 AMQWLWDAGIRACVRRERFHLDSAVYVLSHLERITEEGYVPTAQDVLRSMPPTGI 192
 DB 124 DAIKYLWDPGIGQCYDRRREYQLSDBSTKYLLNDLRDRIATHGYLPTQDDVLRVRVPTGI 183
 QY 193 NEYCFVQKTNLRIVDVGQSKSRKWIHCF:NVIALIYLSLSEYDOCLLENNQENRMK 252
 DB 184 IEYFDLQSVIFRMVDVGQSGSRKWIHCF:NVTSIMFVLSEYDQVLVESDNENRME 243
 QY 253 ESALFGLTILELPWFKSTSVILFNKTDILEEIKPTSHLATYPPSPQPKQDAEAAKRFI 312
 DB 244 ESKALFRTIITYPWFQNSVILFNKDKLLEIKIMYSHLVDPYFPEYDGPORDAQAAAREFI 303
 QY 313 LDMYTRMTGCVDPGSGKKGARSRLFSHYTCATDTQNIKRVKDVRSVLARYLDEIN 372
 DB 304 LKMEVDL-----NPDSK-----IYSHF:CATDTENIRFVFAVKDTIQLNLKEYN 351
 QY 373 LL 374
 DB 352 LV 353

RESULT 5
 GBO_CANFA
 ID GBO_CANFA STANDARD; PRT; 353 AA.
 AC Q28294;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
 GN GNAO.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96433124; PubMed=8836152;
 RA Johnson G.J., Leis L.A., Dunlop P.C.;
 RT "Specificity of G alpha q and G alpha 11 gene expression in platelets
 RT and erythrocytes. Expressions of cellular differentiation and species
 RT differences.";
 RL Biochem. J. 318:1023-1031(1996).
 CC -1- INVOLUTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L76257; AAB05548.1; ALT_INIT.
 DR HSSP; P04896; IAZT.
 DR InterPro; IPR000654; .
 DR InterPro; IPR001019; .
 DR Pfam; PF00503; G-alpha.1.
 DR PRINTS; PR00318; GPROTEINA.
 DR PRINTS; PR00442; GPROTEINAO.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
 KW Palmitate; Lipoprotein.
 FT LIPID 3 3 PALMITATE (BY SIMILARITY).
 FT LIPID 4 4 GTP (BY SIMILARITY).
 FT NP_BIND 40 47 GTP (BY SIMILARITY).
 FT NP_BIND 199 203 GTP (BY SIMILARITY).
 FT NP_BIND 268 271 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT MOD_RES 177 177 (BY SIMILARITY).
 SQ SEQUENCE 353 AA; 41467 MW; E7737307B1F4904C CRC64;

Query Match 53.4%; Score 1047; DB 1; Length 353;
 Best Local Similarity 55.8%; Pred. No. 5.7e-77;
 Matches 202; Conservative 61; Mismatches 87; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDQENRILLOKODRGELKLLGLGSGSGKSTFTKQRIHAGYSE 72
 DB 4 CLSEAEAEARRINDEIERLRDRKROARRELKLLGLTGSGSGKSTFTKQRIHAGYSE 63
 QY 73 EERKGRPLVYQNIYFVSMRAMIEMERLQIPSPRPSKHHASLVMSODPKYVTFEKRYA 132
 DB 64 EDKRGFTKLVQNIYFAMQAMIRAMDTLKIP:KYEHKHAQHLVREVDVEKVSFENPVY 123
 QY 133 AMQWLWDAGIRACVRRERFHLDSAVYVLSHLERITEEGYVPTAQDVLRSMPPTGI 192
 DB 124 DAIKYLWDPGIGQCYDRRREYQLSDBSTKYLLNDLRVADPAYLPTQDDVLRVRVPTGI 183
 QY 193 NEYCFVQKTNLRIVDVGQSKSRKWIHCF:NVIALIYLSLSEYDOCLLENNQENRMK 252
 DB 184 IEYFDLQSVIFRMVDVGQSGSRKWIHCF:NVTSIMFVLSEYDQVLVESDNENRME 243
 QY 253 ESALFGLTILELPWFKSTSVILFNKTDILEEIKPTSHLATYPPSPQPKQDAEAAKRFI 312
 DB 244 ESKALFRTIITYPWFQNSVILFNKDKLLEIKIMYSHLVDPYFPEYDGPORDAQAAAREFI 303
 QY 313 LDMYTRMTGCVDPGSGKKGARSRLFSHYTCATDTQNIKRVKDVRSVLARYLDEIN 372
 DB 304 LKMEVDL-----NPDSK-----IYSHF:CATDTENIRFVFAVKDTIQLNLKEYN 351
 QY 373 LL 374
 DB 352 LV 353

RESULT 6
 GBO_RAT
 ID GBO_RAT STANDARD; PRT; 353 AA.
 AC P82471;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
 GN GNAO.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;

QY 133 AMQWLWRDAGIRACVYERREFFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPPTGI 192
 Db 130 SAKITLWDPGIEQYDRREYQLSDSAKYLTLDVDRATSGYLTQDQDVLRSRMPPTGI 189
 QY 193 NEYCFVQKTNLRIVDVGQOKSERKKWIKHCFENFVIALIYLSLSEYDQCLLENNOENRMK 252
 Db 190 IEYFPDLNFIIFRMDVVGQORSEKRWIKHCFENFVIALIYLSLSEYDQCLLENNOENRMK 249
 QY 253 ESALFGLTELEPWFKSTSVILFNKTDILEEIKPTSHLATYFPFSGPKODAAAKRFI 312
 Db 250 ESKALFTVITYFPWFQNSVILFNKTDILEEIKPTSHLATYFPFSGPKODAAAKRFI 309
 QY 313 LDMYTRMTGCVDPGSGKAGSRRLFSHYTCATDTONIRKVKDVRDSVLARYLDEIN 372
 Db 310 LKMFVDL-----NPDSDK-----IIYSHFTCATDTONIRFVFAAVKDTILOLNKEYN 357
 QY 373 LL 374
 Db 358 LV 359

RESULT 8
 ID GBQ_MOUSE STANDARD; PRT: 353 AA.
 AC P21279;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
 GN GNAO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=91067657; PubMed=2123549;
 RA Strathmann M., Simon M.I.;
 RT "G protein diversity: a distinct class of alpha subunits is present
 in vertebrates and invertebrates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9113-9117(1990).
 RN [2]
 RP PALMITOYLATION.
 RC TISSUE=Brain;
 RX MEDLINE=94043367; PubMed=8227083;
 RA Wedegaertner P.B., Chu D.H., Wilson P.T., Lewis M.J., Bourne H.R.;
 RT "Palmitoylation is required for signaling functions and membrane
 attachment of Gq alpha and Gs alpha.";
 RL J. Biol. Chem. 268:25001-25008(1993).
 CC -|- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 SIGNALING SYSTEMS.
 CC -|- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -|- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M55412; AAA63306.1; ALT_INIT.
 DR FIR; A38414; RGMQO.
 DR HSP; P04696; IAZT.
 DR MGD; MGI:95776; Gnaq.
 DR InterPro; IPR000654; -.
 DR InterPro; IPR001019; -.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.

DR PRINTS; PR00442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
 KW Palmitate; Lipoprotein.
 FT LIPID 3 PALMITATE.
 FT LIPID 4 4
 FT NP_BIND 40 47 GTP (BY SIMILARITY).
 FT NP_BIND 199 203 GTP (BY SIMILARITY).
 FT NP_BIND 268 271 GTP (BY SIMILARITY).
 FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT (BY SIMILARITY).
 SQ SEQUENCE 353 AA; 41478 MW; C41B2AC11C674C5F CRC64;

Query Match 53.3%; Score 1344; DB 1; Length 353;
 Best Local Similarity 55.5%; Pred. No. 1e-76; 87; Indels 12; Gaps 2;
 Matches 201; Conservative 62; Mismatches 87;

QY 13 CLTEDEAAARVDQOEINRILLOKQKODRGELKLLGLPGESGKSTFIKQRIIHGAGYSE 72
 Db 4 CLSEAEKARRINDEIERHVRDKRDARRELKLLGLTGESGKSTFIKQRIIHGAGYSD 63
 QY 73 EERKGRPLVYONIFVSMRAMIEAMERLQIPFSRPESKHHSALVMSODPYKVTTFEKRYA 132
 Db 64 EDKRGFTKLVIYONIFVSMRAMIEAMERLQIPFSRPESKHHSALVMSODPYKVTTFEKRYA 123
 QY 133 AMQWLWRDAGIRACVYERREFFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPPTGI 192
 Db 124 DAISLWNPDPGIEQYDRREYQLSDSAKYLTLDVDRATSGYLTQDQDVLRSRMPPTGI 183
 QY 193 NEYCFVQKTNLRIVDVGQOKSERKKWIKHCFENFVIALIYLSLSEYDQCLLENNOENRMK 252
 Db 184 IEYFPDLNFIIFRMDVVGQORSEKRWIKHCFENFVIALIYLSLSEYDQCLLENNOENRMK 243
 QY 253 ESALFGLTELEPWFKSTSVILFNKTDILEEIKPTSHLATYFPFSGPKODAAAKRFI 312
 Db 244 ESKALFTVITYFPWFQNSVILFNKTDILEEIKPTSHLATYFPFSGPKODAAAKRFI 303
 QY 313 LDMYTRMTGCVDPGSGKAGSRRLFSHYTCATDTONIRKVKDVRDSVLARYLDEIN 372
 Db 304 LKMFVDL-----NPDSDK-----IIYSHFTCATDTONIRFVFAAVKDTILOLNKEYN 351
 QY 373 LL 374
 Db 352 LV 353

RESULT 9
 ID GBQ_HUMAN STANDARD; PRT: 353 AA.
 AC P50148; Q13462; Q92471; O15108;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
 GN GNAO OR GAO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96423032; PubMed=8825633;
 RA Dong Q., Shenker A., Way J., Haddad B.R., Lin K., Hughes M.R.,
 McBride W.O., Spiegel A.M., Batty J.;
 RT "Molecular cloning of human G alpha q cDNA and chromosomal
 localization of the G alpha q gene (GNAO) and a processed
 pseudogene.";
 RL Genomics 30:470-475(1995).
 RN [2]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Prostate;
 RX MEDLINE=96256639; PubMed=8664309;
 RA Chen B., Leverette R.D., Schwinn D.A., Kwatra M.M.;

RC TISSUE-Hematopoietic;
RX MEDLINE=96077138; PubMed=7492305;
RA Thomas C.P., Dunn M.J., Mattera R.;
RT "Ca2+ signalling in K562 human erythroleukaemia cells: effect of
RT dimethyl sulphoxide and role of G-proteins in thrombin- and
RL thromboxane A2-activated pathways.";
RL Biochem. J. 312:151-158(1995).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M69013; AAC58624.1;
CC EMBL: AF011497; AAB64303.1;
CC EMBL: AC005262; AAC25615.1;
CC EMBL: L40630; AAC99949.1;
CC PIR: A39394; RCHUY.
CC HSSP: P04896; LAZT.
CC MIM: 139313;
CC InterPro: IPR000654;
CC InterPro: IPR001019;
CC Pfam: PF00503; G-alpha.1;
CC PRINTS: PR00318; GPROTEINA.
CC PRINTS: PR00442; GPROTEINAQ.
CC GTP-binding; transducer; Multigene family; ADP-ribosylation.
CC NP_BIND 46 53 GTP (BY SIMILARITY).
CC NP_BIND 205 209 GTP (BY SIMILARITY).
CC NP_BIND 274 277 GTP (BY SIMILARITY).
CC MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)
CC (BY SIMILARITY).
CC CONFLICT 6 6 M -> I (IN REF. 2).
CC CONFLICT 266 266 H -> H (IN REF. 4).
CC CONFLICT 285 285 Y -> H (IN REF. 4).
CC CONFLICT 301 302 DA -> EP (IN REF. 1).
CC CONFLICT 310 310 L -> P (IN REF. 2).
CC SEQUENCE 359 AA; 42123 MW; DD37176589E66046 CRC64;

Query Match 53.1%; Score 1040; DB 1; Length 359;
Best Local Similarity 55.8%; Pred. No. 2.1e-76;
Matches 202; Conservative 65; Mismatches 83; Indels 12; Gaps 2;
QY 13 CLTEDEKAARVDQENRILLBOKKODRGELKLLILGPGESGKSTFIKQMRILHAGYSE 72
DB 10 CLSDEVKESKRINAETEKQLRRDKRDARRELKLLILGTSGSKSTFIKQMRILHAGYSE 69
QY 73 EERKGFRLPLVQNIYFVMSRAMEERLQIPSRPESKHHSLVMSODPKYVTTTEKRYA 132
DB 70 EDKRGFTKLVIQNIITAMQAMIRAMETLKLKYEQNKANALLIREVDVKEVITFEHQIV 129
QY 133 AAMQNLWRDAGTRACYERREPHILDSAVYVLSHLERITEEGYVPTAQDVLRSRMTTGI 192
DB 130 SAIKLTWEDPGIQECYDRREYQLSDSAKYILTDVDRATLGLTLPQDVLKRVPTTGI 189
QY 193 NEYCFVQKTNLRIVYVGQSKERKWTCHCFENVIALYLSLSEYDQCLENNQENRMK 252
DB 190 IYFPFDLENIIFRWYDVGGQSRERKWHCFENFVTSIMPLVALSEYDQVLYESDNRME 249
QY 253 ESLALFGTILELPWPKSTSVILFLNKTIDILEKIPSTSHLATYFSPFGQPKDAEAKRFI 312
DB 250 ESKALFRITIIYFPQNSVILFLNKKDLEKILYSLVDFPFEDFGPQDQAAREFI 309
QY 313 LDMYTRMYTCVGDGEGSKGARSRLFSHYTCATDPTQNIKKVFKVDVRDVLARYLDEIN 372

Db 310 LKMEVDL-----NPDSDK-----IIYSHFTCATDTENIRFVFAAVKDTIOLNLKEYN 357
QY 373 LL 374
Db 358 LV 359

RESULT 11
ID GBLL_MOUSE STANDARD; PRT; 359 AA.
AC P21278; Q661939;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN, 'ALPHA-11' SUBUNIT.
GN GNALL OR GNA-11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91067657; PubMed=2123549;
RA Strathmann M., Simon M.I.;
RT "G protein diversity: a distinct class of alpha subunits is present
RT in vertebrates and invertebrates.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9111-9117(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=129/SV;
RC MEDLINE=96435432; PubMed=8838318;
RA Davignon I., Barnard M., Gavrilova O., Sweet K., Wilkie T.M.;
RT "Gene structure of murine Gna11 and Gna15: tandemly duplicated Gq
RT class G protein alpha subunit genes.";
RL Genomics 31:359-366(1996).
RN [3]
RP SEQUENCE OF 211-271 FROM N.A.
RX MEDLINE=90017488; PubMed=2508088;
RA Strathmann M., Wilkie T.M., Simon M.I.;
RT "Diversity of the G-protein family: sequences from five additional
RT alpha subunits in the mouse.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7407-7409(1989).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M55411; AAA63305.1;
CC EMBL: U37413; AAB36839.1;
CC EMBL: U37411; AAB36839.1; JOINED.
CC EMBL: U37412; AAB36839.1; JOINED.
CC EMBL: M57617; AAA63301.1;
CC PIR: B38414; RGMS11.
CC PIR: B38833; B38833.
CC HSSP: P04896; LAZT.
CC MGD: MGI:95766; Gna11.
CC InterPro: IPR000654;
CC InterPro: IPR001019;
CC Pfam: PF00503; G-alpha.1.
CC PRINTS: PR00318; GPROTEINA.
CC PRINTS: PR00442; GPROTEINAQ.
CC GTP-binding; Transducer; Multigene family; ADP-ribosylation.
KW

Thu Sep 6 15:16:58 2001

CC	use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; U89139; AAB49314.1; -
DR	InterPro; IPR000654; -
DR	InterPro; IPR001019; -
DR	Pfam; PF00503; G-alpha: 1.
DR	PRINTS; PR00318; GPROTEINA.
DR	PRINTS; PR00442; GPROTEINAQ.
DR	GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW	Palmitate; Lipoprotein.
KW	LIPID 3 3 PALMITATE (BY SIMILARITY).
FT	LIPID 4 4 GTP (BY SIMILARITY).
FT	NP-BIND 40 47 GTP (BY SIMILARITY).
FT	NP-BIND 199 203 GTP (BY SIMILARITY).
FT	NP-BIND 268 271 GTP (BY SIMILARITY).
FT	MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT	(BY SIMILARITY).
SQ	SEQUENCE 353 AA; 41521 MW; 218B92CA16540408 CRC64;
Query Match 52.9%; Score 1036; DB 1; Length 353;	
Best Local Similarity 56.4%; Pred. No. 4.4e-76;	
Matches 204; Conservative 59; Mismatches 87; Indels 12; Gaps 2;	
QY	13 CLTEDEAAARVDQEIINRILLEQKQKQDRELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72
DB	4 CLSBEAKQKRNQIEIQRKRDARRELKLLLLGTGSGKSTFIKQMRIIHGAGYSD 63
QY	73 BERKGRPLVYQNIQFVSMRAMIEMERLQIPFSPKSHASLVMSQDPYKVTTFEKR 132
DB	64 EDKRGFKLVQNFQIFMAMQSMIRMDLIQISYGDANSIEHADLVRSVDYESTVTFEEPY 123
QY	133 AAMWLWDAGIRACRYERRERFHLDSAVYVLSHLERITEEGYVPTAQDVLRSRPTTGI 192
DB	124 TAMNSLWQDTGIQHCYDRREYQLTDSAKYLLTDLRIAADKYVSTLQDILRVAPTTGI 183
QY	193 NEYCFVQKTNLRIVDVGQSKSRKWIHCFENVALIYLSLEYDQCLLENNOENRMK 252
DB	184 IEYFDLEIRFRWVDVGQSRERKWIHCFENVTIIFLVALSEYDQILFESONENRME 243
QY	253 ESALFGTILELPWFKSTSVILFNKTLDEKIPKTSHTLATYFPSPGPKQDAEAKRFI 312
DB	244 ESKALFKTIITYPWFQHSVILFNKDLLEKIMVSHLVDFPEYDGPDKDAIAAREFI 303
QY	313 LDMYTRMYTCVDPGSGKSGARSRLFSHYTCATDTQNIKRVKDVDRSVLARYLDEIN 372
DB	304 LRMFVEL-----NPDPEK-----IIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 351
QY	373 LL 374
DB	352 LV 353
RESULT 13	
GB11_MELGA	STANDARD; PRT; 359 AA.
ID	GB11_MELGA
AC	P45645;
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-11 SUBUNIT.
GN	GNAL1.
OS	Meleagris gallopavo (Common turkey).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
OX	NCBI_TaxID=9103;
RN	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC	TISSUE=Brain;
RX	MEDLINE-93207527; PubMed-8457205;

RA Maurice D.H., Waldo G.L., Morris A.J., Nicholas R.A., Harden T.K.;
 RT "Identification of G alpha 11 as the phospholipase C-activating G-
 RL Biochem. J. 290:765-770(1993)."
 CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
 CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X73072; CAA51530.1; -
 CC HSSP; P04896; IAZT; -
 CC InterPro; IPR000654; -
 CC InterPro; IPR001019; -
 CC Pfam; PF00503; G-alpha.1;
 CC PRINTS; PR00318; GPROTEINA.
 CC PRINTS; PR00442; GPROTEINAO.
 CC GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 CC NP-BIND; 46 53 GTP (BY SIMILARITY).
 CC FT NP-BIND 205 209 GTP (BY SIMILARITY).
 CC FT NP-BIND 274 277 GTP (BY SIMILARITY).
 CC FT MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 CC FT (BY SIMILARITY).
 CC SEQUENCE 359 AA; 42058 MW; B9E16427169BE1AE CRC64;

 Query Match 52.8%; Score 1035; DB 1; Length 359;
 Best Local Similarity 55.5%; Pred. No. 5.4e-76;
 Matches 201; Conservative 65; Mismatches 84; Indels 12; Gaps 2;
 QY 13 CLTEDEKAAARVQOEINRILLEOKKQDRGELKLLLLGPGSGKSTFKOMRIHGCYSE 72
 Db 10 CLSDEVKESKRINAELEKQLRRDKRDARRELKLLLLGTGESGKSTFKOMRIHGCYSE 69
 QY 73 EERKGFPLVYQNIFFVSMRAMIEMERLQIPF:RPESKHHASLVMSODPYKVTTEKRYA 132
 Db 70 EDKKGFTKLVYQNIFFVSMRAMIEMERLQIPF:RPESKHHASLVMSODPYKVTTEKRYA 132
 QY 133 AAMQWLWDAGIRACVYERREHLLDSAVYLLSHLERITEEGYVPTAQDVLRSMPPTGI 192
 Db 130 SAKTLWDPGQECYDRREYQSDSAKYLLSDVDRIATVGYLPTQDDVLRVVRPTGI 189
 QY 193 NEYCFVQKTNLRIVDVGGOKSERKKWICHFENVIALLYLASLSEYDQCLEENNRMK 252
 Db 190 IEYFPDLENIIFRMVDVGGQSRERRKWIHCFFENVTIMFVLSSEYDQVLVESDNENRME 249
 QY 253 ESLALFCTILELPWFKSTSVILFLNKTDLILEEKIPTSHLATYFPSPQGPQDAEAAKRFI 312
 Db 250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEDKILYSHLVVDVFPFDPGPDQDAQAAREFI 309
 QY 313 LDMYTRMYTCGVDGPGEGSKGARSRLFSHYTCATDTONIRKVKDVRDVSILARYLDETN 372
 Db 310 LKMFVDL-----NPDSDK-----IIYSHFTCATDNTENIREVFAVKDTILQLNLKYN 357
 QY 373 LL 374
 Db 358 LV 359
 RESULT 14
 GB11_RAT
 ID GB11_RAT
 AC Q9JID2;
 DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-11 SUBUNIT.
 GN GN11.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Crania:a; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciuroinathi; Muridae; Murinae; Rattus.
 CC NCBI_Taxid-10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RA Strotmann R.;
 RT "Rattus norvegicus guanine nucleotid; binding protein alpha 11 subunit
 RT (G11).";
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF239674; AAP81690.1; -
 CC InterPro; IPR000654; -
 CC InterPro; IPR001019; -
 CC PRINTS; PR00318; GPROTEINA.
 CC PRINTS; PR00442; GPROTEINAO.
 CC GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 CC NP-BIND; 46 53 GTP (BY SIMILARITY).
 CC FT NP-BIND 205 209 GTP (BY SIMILARITY).
 CC FT NP-BIND 274 277 GTP (BY SIMILARITY).
 CC FT MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 CC FT (BY SIMILARITY).
 CC SEQUENCE 359 AA; 42026 MW; B4CD057E9FC7092A CRC64;

 Query Match 52.6%; Score 1030; DB 1; Length 359;
 Best Local Similarity 55.0%; Pred. No. 1.4e-75;
 Matches 199; Conservative 66; Mismatches 85; Indels 12; Gaps 2;
 QY 13 CLTEDEKAAARVQOEINRILLEOKKQDRGELKLLLLGPGSGKSTFKOMRIHGCYSE 72
 Db 10 CLSDEVKESKRINAELEKQLRRDKRDARRELKLLLLGTGESGKSTFKOMRIHGCYSE 69
 QY 73 EERKGFPLVYQNIFFVSMRAMIEMERLQIPF:RPESKHHASLVMSODPYKVTTEKRYA 132
 Db 70 EDKKGFTKLVYQNIFFVSMRAMIEMERLQIPF:RPESKHHASLVMSODPYKVTTEKRYA 132
 QY 133 AAMQWLWDAGIRACVYERREHLLDSAVYLLSHLERITEEGYVPTAQDVLRSMPPTGI 192
 Db 130 SAKTLWDPGQECYDRREYQSDSAKYLLSDVDRIATVGYLPTQDDVLRVVRPTGI 189
 QY 193 NEYCFVQKTNLRIVDVGGOKSERKKWICHFENVIALLYLASLSEYDQCLEENNRMK 252
 Db 190 IEYFPDLENIIFRMVDVGGQSRERRKWIHCFFENVTIMFVLSSEYDQVLVESDNENRME 249
 QY 253 ESLALFCTILELPWFKSTSVILFLNKTDLILEEKIPTSHLATYFPSPQGPQDAEAAKRFI 312
 Db 250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEDKILYSHLVVDVFPFDPGPDQDAQAAREFI 309
 QY 313 LDMYTRMYTCGVDGPGEGSKGARSRLFSHYTCATDTONIRKVKDVRDVSILARYLDETN 372
 Db 310 LKMFVDL-----NPDSDK-----IIYSHFTCATDNTENIREVFAVKDTILQLNLKYN 357
 QY 373 LL 374

QY 121 PYKVTTFKRYAAAMQWLWRDAGIRACYERRRERHLLDSAVYLSHLERITEEGVPTAQ 180
DB 118 VEKVCCTFEQYVNAIKNLWSDGICQEDRRREYQLSDSTKYLLTDVDRISKPGYLPQQ 177
QY 181 DVLRSRMPPTTGINEYCFVSKTNLRIVDVGOKSKRWKHCFFENVIALIYLASEYDQ 240
DB 178 DVLRLVRPPTGIIIEYFDFLENIIFRMVDVGGQSRERKRWKHCFFENVTSIMFLVALSEYDQ 237
QY 241 CLEENNOENRKESLALFCTILELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSG 300
DB 238 VLVESDNENRMEESKALFRTIITYPFQNSVILFNKDKLLEDKIMYSLVDYPPFEDG 297
QY 301 PKODAAAKRFRILDMYTRMYTCVCDGPEGSKGARSRRFLFSHYTCATQNTONRKVKDYR 360
DB 298 PQDAATAREFLLKMFVDL-----NPDSK-----IIVSHFTCATDTENIRFVFAAYK 345
QY 361 DSVIARYLDEINLL 374
DB 346 DTILOHNLKEYNLV 359

Search completed: September 6, 2001, 10:55:23
Job time: 103 sec

Db 358 LV 359

RESULT 15
GB11_XENLA STANDARD; PRT: 359 AA.
AC P43444; 1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-II SUBUNIT.
GN Xn11.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=94298961; PubMed=8026589;
RA Shapira H., Way J., Lipinsky D., Oron Y., Battay J.F.;
RT "Neuromedin B receptor, expressed in Xenopus laevis oocytes,
selectively couples to G alpha q and not G alpha 11.";
RL FEBS Lett. 348:89-92(1994).
RN [2]
RP ERRATUM
RA Shapira H., Way J., Lipinsky D., Oron Y., Battay J.F.;
RL FEBS Lett. 349:318-318(1994).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U10494; AAA52188.1; -
CC HSSP; P04896; 1A2T.
CC InterPro: IPR000654; -
CC InterPro: IPR001019; -
CC Pfam: PF00503; G-alpha: 1.
CC PRINTS; PR00318; GPROTEINA.
CC PRINTS; PR00442; GPROTEINAQ.
CC GTP-binding; Transducer; Multigene family; ADP-ribosylation.
KW NP_BIND 46 53 GTP (BY SIMILARITY).
FT NP_BIND 205 209 GTP (BY SIMILARITY).
FT NP_BIND 274 277 GTP (BY SIMILARITY).
FT MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX) (BY
SIMILARITY).
FT SEQUENCE 359 AA; 42088 MW; DA6B376993FDD870 CRC64;

Query Match 52.4%; Score 1026.5; DB 1; Length 359;
Best Local Similarity 54.0%; Pred. No. 2.6e-75;
Matches 202; Conservative 65; Mismatches 92; Indels 15; Gaps 3;
QY 1 MARSITWPCPCWCTEDEKAARVDQENRILLQKQDRGCEKLLLLGPGESCKSTFIK 60
DB 1 MTLDSTMAC---CLSEEVKSKRNAEIEKQLRRDKDSRRELKLLLLGTGSGKSTFIK 57
QY 61 QMRILHAGYSEERKFGRLPYQNIYFVMSRAMTEAMERLQIPFSPRESKHHASLVMSQD 120
DB 58 QMRILHAGYSEERKFGRLPYQNIYFVMSRAMTEAMERLQIPFSPRESKHHASLVMSQD 117

Thu Sep 6 15:16:58 2001

us-08-878-801-2.rsp

Page 12

Q9hfn1 blumeria gr
Q90847 gallus gall
O13055 oryzias lat
Q9uqa4 homo sapien
Q18205 caenorhabdi
Q9xav3 geodia cydo
Q9n193 octopus vul
Q90846 gallus gall
Q9y201 ephydatia f
P79895 sparus aca
Q9dg28 gallus gall
Q9n194 octopus vul
Q9y123 ambystoma t
Q9y122 ambystoma t
O93565 xenopus lae
O13011 xenopus lae
Q9vs04 drosophila
Q9dg27 gallus gall
Q22567 caenorhabdi
Q00581 cryphonectr
O74204 podospora a
Q93564 xenopus lae
Q9hfw7 neurospora
Q94216 sclerotinia
Q9hfa3 rosellinia

20 745.5 38.0 354 3 Q9HFN1
21 744.5 38.0 354 13 Q90847
22 742 37.9 355 13 O13055
23 741.5 37.8 354 4 Q9UGA4
24 741.5 37.8 354 5 Q18205
25 738 37.7 359 5 Q9XAV3
26 734.5 37.5 354 5 Q9N193
27 732.5 37.4 354 13 Q90846
28 732 37.3 301 5 Q9Y201
29 728 37.1 355 13 Q9W6A4
30 727 37.1 350 13 P79895
31 726.5 37.1 350 13 Q9DG28
32 725.5 37.0 354 5 Q9N194
33 722.5 36.9 350 13 Q9Y123
34 718 36.6 354 13 Q9Y122
35 714 36.4 350 13 Q93565
36 713.5 36.4 354 13 O13011
37 713.5 36.4 355 5 Q9VS04
38 712.5 36.4 354 13 Q9DG27
39 705 36.0 357 5 Q22567
40 700 35.7 355 3 Q00581
41 696.5 35.5 354 3 O74204
42 695.5 35.5 354 13 Q93564
43 695.5 35.5 356 3 Q9HFW7
44 694 35.4 355 3 Q94216
45 686 35.0 356 3 Q9HFA3

ALIGNMENTS

RESULT 1
Q9TU29 PRELIMINARY: PRT; 374 AA.
ID Q9TU29
AC Q9TU29;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)
DE GALPHAL6.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20035837; PubMed=10571060;
RA Feild J.A., Foley J.J., Testa T.F., Nuthulaganti P., Ellis C.,
RA Sarau H.M., Ames R.S.;
RT "Cloning and characterization of a rabbit ortholog of human Galphal6
and mouse Galphal5";
RL FEBS Lett. 460:53-56(1999).
DR EMBL; AF169627; AAF06740.1; .
DR HSSP; P10824; IAS3.
DR InterPro; IPR001019; .
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
SQ SEQUENCE 374 AA; 42933 MW; 630B2920A208CAIE CRC64;

Query Match 87.3%; Score 1711; DB 6; Length 374;
Best Local Similarity 85.8%; pred. No. 9.4e-134; Indels 0; Gaps 0;
Matches 321; Conservative 29; Mismatches 24;

QY 1 MARSLTWRCPCWCLTEDEKAAARVDQENRILLEEKKQDRGELKLLKLLGPESGKSTIK 60
Db 1 MARSLAWRCPCWCLTEDEKAAARVDQENRILLEEKKQDRGELKLLKLLGTSKGSTIK 60
QY 61 QMRITHGAGYSEERKGFPLVYQNIYFSMRAMIEMERLQIPFSPESKHASLVMSOD 120
Db 61 QMRITHGAGYSEERKGFPLVYQNIYFSMRAMIEMERLQIPFSPESKHASLVMSOD 120
QY 121 PYKVTTTEFKRYAAMQWLWDACIRACYERRREFHLLDSAVVYLSHLERITEEGVYPTAQ 180

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: September 6, 2001, 10:53:40 ; Search time 24.8 Seconds
(without alignments)
1995.245 Million cell updates/sec

Title: US-08-878-801-2
Perfect score: 1960
Sequence: 1 MARSLTWRCPCWCLTEDEKA.....VFKDVRDSVLARYLDEINLL 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL16:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organellie:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_unclassified:
13: sp_vertebrate:
14: sp_virus:

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	1711	87.3	374	6	Q9TU29	Q9tu29 oryctolagus
2	1031	52.6	353	5	Q9U473	Q9u473 panulirus a
3	1026	52.3	353	5	Q9V6E2	Q9v6e2 drosophila
4	1006	51.3	353	5	Q9N192	Q9n192 octopus vul
5	1001	51.1	353	5	Q9N192	Q9n192 calliphora
6	986	50.3	353	5	P91955	P91955 limulus pol
7	974	49.7	355	5	Q17386	Q17386 caenorhabdi
8	876.5	44.7	355	5	Q9XZV4	Q9xzv4 geodia cydo
9	829.5	42.3	305	5	Q9Y207	Q9y207 hydra magni
10	790	40.3	279	5	Q917C8	Q917c8 drosophila
11	787.5	40.2	334	3	Q9HFE9	Q9hf99 schizophyll
12	764.5	39.0	353	3	Q9HFA2	Q9hfa2 rosellinia
13	755.5	38.5	353	3	Q9HEP9	Q9hep9 botrytis ci
14	754	38.5	357	5	Q9N197	Q9n197 halocynthia
15	753.5	38.4	353	3	P78705	P78705 neurospora
16	752.5	38.4	352	5	Q21517	Q21917 caenorhabdi
17	751	38.3	305	5	Q9Y203	Q9y203 ephydatia f
18	746.5	38.1	351	5	Q9Y202	Q9y202 ephydatia f
19	746	38.1	357	5	Q9N2V6	Q9n2v6 caenorhabdi

```

Db 121 PYKNTETRYALAVQSLWDAGVACRYERRERFHLDSAVYLSHLRTAEYGYPTAQ 180
QY 181 DVLSRMTPTGINEXCFVQKTNLVDVGGQKSKRWKWHCFENVIALYLSASEYDQ 240
Db 181 DVLSRMTPTGINEXCFVQKTNLVDVGGQKSKRWKWHCFENVIALYLSASEYDQ 240
QY 241 CLEENQENRMKESALFETLLEPWFKSTVTLFUNKTDILEEKIPTSHLATYFPFSG 300
Db 241 CLEENQENRMKESALFETLLEPWFKSTVTLFUNKTDILEEKIPTSHLATYFPFSG 300
QY 301 PKQDAEAKRFDLDMRYMTGCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKDV 360
Db 301 PKQDAEAKRFDLDMRYMTGCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKDV 360
QY 361 DSVLYRLDEINLL 374
Db 361 DSVLYRLDEINLL 374

RESULT 2
Q90473 PRELIMINARY; PRT; 353 AA.
AC Q90473
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE G-ALPHA-49B OR CGI17759.
OS Panulirus argus (Spiny lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;
OC Palinuroidea; Palinuridae; Panuliridae; Panulirus.
OX NCBI_TaxID=6737;
RN [1]
RP SEQUENCE FROM N.A.
RC TSSUB-OLFACTORY ORGAN;
RA Mungar S.D., Gleeson R.A., Aldrich H.C., Rust N.C., Ache B.W.,
RA Greenberg R.M.;
RT "Molecular Evidence for Phosphoinositide-Mediated Signaling in Lobster
RT Olfactory Receptor Neurons."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF201328; AAF19378.1;
DR HSP: P10824; 180F.
DR InterPro: IPRO01019;
DR DR: IPRO01019;
DR DR: PF00503; G-alpha; 1.
DR PRINTS: PR00318; GPROTEIN.
DR PRODOM: PD001018; 1.
DR SMART: SM00275; G-alpha; 1.
SQ SEQUENCE 353 AA; 41478 MW; FD9B551F66327BAA CRC64;

Query Match 52.6%; Score 1031; DB 5; Length 353;
Best Local Similarity 56.1%; Pred. No. 1.8e-77;
Matches 203; Conservative 59; Mismatches 88; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDQETNRIILLOKQDRGELKLLLLGPGSGKSTFIKQMRILHAGYSE 72
Db 4 CLSEAKEQKRNQETNRIILLOKQDRGELKLLLLGPGSGKSTFIKQMRILHAGYSD 63
QY 73 EERKGFRLVLYONIFVSRMATEAMERLQIPSPESKHASLVMSQDPYKVTTFERYA 132
Db 64 EDRKGFRLVLYONIFVSRMATEAMERLQIPSPESKHASLVMSQDPYKVTTFERYA 123
QY 133 AMOWLWDAGTRACYERRERFHLDSAVYLSHLRTAEYGYPTAQDVLRSRMTPTGI 192
Db 124 TAMKSLWQDTGIQHCYDRREYQITDSAKYLLDLDRIAATDYSTIQDILRVRAPTGI 183
QY 193 NEYCFVQKTNLVDVGGQKSKRWKWHCFENVIALYLSASEYDQCLEENQENRMK 252
Db 184 IEXFPDLEIEFRVMDVGGQKSKRWKWHCFENVIALYLSASEYDQILEENQENRM 243
QY 253 ESLALFTGILPWFKSTVTLFUNKTDILEEKIPTSHLATYFPFSGQPKQDAEAAKRFI 312

```

```

Db 244 ESKALFKTILITYPWFQSSVILFNKKDLLEKIMYSHLVDFPEYDGPDKRDAAREFI 303
QY 313 LDMYTRMTGCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKDVRSVLYRLDEIN 372
Db 304 LRMFVEL-----NPDEK-----IYSHF-CATDTENIRFVFAAVKDTILQLNLKEYN 351
QY 373 LL 374
Db 352 LV 353

RESULT 3
Q9V6E2 PRELIMINARY; PRT; 353 AA.
AC Q9V6E2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE G-ALPHA-49B OR CGI17759.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
RA Bozina D., Botchan M.R., Bouck J., Bockstein P., Brothier P.,
RA Butts J.M., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stاپleton M., Strong R., Sun E.,
RA Swirskas R., Taatjes R., Turner K., Turner K., Wang A.H., Wang X.,
RA Wang 2-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Verter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AF003821; AAF58485.2;
DR HSP: P10824; 180F.
DR Flybase: FBgn0004435; G-alpha-49B.
DR InterPro: IPR001019;
DR Pfam: PF00503; G-alpha; 2.

```

1

Db 184 LEYFPLDGLGIVFMVDGQGRERRKWHICFENVTSIIIFVLSAYDQILFESDNENRME 243
 QY 253 ESALFGLTILELPWPKSTSVILFNKTDILEEKIPTSHLYTFPSFGQKODAAAKRFI 312
 Db 244 ESALFRTIITYPWONSSVILFNKDLLEKIMYSHLVDFYDFYDGPCKDHAARQFV 303
 QY 313 LDMYTRMTGCVGDEGSGKKGARSRLSHYTCATDTONIRKVFKDVRSDSVLARYLDEIN 372
 Db 304 L-----KKYLAANPDPE-----RQCYSHFTTATDTEINIKLVCAVKDTIMQNALKEFN 351
 QY 373 L 373
 Db 352 L 352
 RESULT 6
 ID P91955 PRELIMINARY; PRT: 353 AA.
 AC P91955;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DE GQ PROTEIN ALPHA SUBUNIT.
 OS Limulus polyphemus (Atlantic horseshoe crab).
 OC Eukaryota; Metazoa; Chelicerata; Merostomata; Xiphosura;
 OC Limulidae; Limulus.
 OX NCBI_TaxID=6850;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=VENTRAL EYE;
 RA Mungier S.D., Schrenser-Berlin J.L., Brink C.M., Battelle B.A.;
 RA Invert. Neurosci. 0:0-0(0).
 DR EMBL; U88586; AAB48510.1; -.
 DR HSP; P10824; IBOF.
 DR InterPro; IPR001019; -.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR SMART; SM00275; G-alpha; 1.
 SQ SEQUENCE 353 AA; 41497 MW; C1C57783B3D2D516 CRC64;

Query Match 50.3%; Score 986; DB 5; Length 353;
 Best Local Similarity 53.9%; Pred. No. 9.5e-74;
 Matches 195; Conservative 62; Mismatches 93; Indels 12; Gaps 2;
 QY 13 CLTEDEKAARVDQENRILLEOKKODRGELKLLILPGESGKSTFIKQRIIHHGAGYSE 72
 Db 4 CLSEEGKEQRINQETEROLRDKRDARRELKLLILGTGSGKSTFIKQRIIHHGAGYSD 63
 QY 73 EERKGRPLVYQNIYFMSRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 132
 Db 64 DKKSYIKLVYQNIYFMSRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 132
 QY 133 AMQWLWDAGIRACVYERRREFHLLDSAVYVLSHLERITEEGYVPTAQDVLRSRMPPTGI 192
 Db 124 EAKLSLVDPGIGECYDRREYOLDSAKYLLINDIDRIVPNYLPQQDILKRVVRPTTGI 183
 QY 193 NPYCSVQKTNRIYVDVGQSKERRKWHICFENVTSIIIFVLSAYDQILFESDNENRME 252
 Db 184 IYFPLDLSIIIFMVDVGQSKERRKWHICFENVTSIIIFVLSAYDQILFESDNENRME 243
 QY 253 ESALFGLTILELPWPKSTSVILFNKTDILEEKIPTSHLYTFPSFGQKODAAAKRFI 312
 Db 244 ESALFRTIITYPWONSSVILFNKDLLEKIMYSHLVDFYDFYDGPCKDHAARQFV 303
 QY 313 LDMYTRMTGCVGDEGSGKKGARSRLSHYTCATDTONIRKVFKDVRSDSVLARYLDEIN 372
 Db 304 LKMFVDL-----NPDSEK-----IYSHFTCATDTEINIRFVFAVKDITLQLNLKEYN 351
 QY 373 L 374
 Db 352 LV 353

RESULT 7
 ID Q17386 PRELIMINARY; PRT: 355 AA.
 AC Q17386; 002546;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE G1-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE EGL-30.
 GN EGL-30 OR M01D7.7
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2;
 RC MEDLINE=96221161; PubMed=8630258;
 RA Brundage L., Avery L., Katz A., Kim U.J., Mendel J.E., Sternberg P.W.,
 RA Simon M.I.;
 RT "Mutations in a C. elegans Gqalpha gene disrupt movement, egg laying,
 RT and viability.";
 RL Neuron 16:999-1009(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RC MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Gpsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Lalister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roop's A., Saunders D., Shownkeen R.,
 RA Smaildon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
 RA Watson A., Wellstock L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Gattung S., Goela D., Wilson R.;
 RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U56864; AAB04059.1; -.
 DR EMBL; AF003739; AAB58071.1; -.
 DR HSP; P10824; IBOF.
 DR InterPro; IPR001019; -.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR SMART; SM00275; G-alpha; 1.
 SQ SEQUENCE 355 AA; 41865 MW; FE438B01C2E1355C CRC64;
 Query Match 49.7%; Score 974; DB 5; Length 355;
 Best Local Similarity 53.0%; Pred. No. 9.4e-73;
 Matches 193; Conservative 64; Mismatches 93; Indels 14; Gaps 3;
 QY 13 CLTEDEKAARVDQENRILLEOKKODRGELKLLILPGESGKSTFIKQRIIHHGAGYSE 72
 Db 4 CLSEEGKEQRINQETEROLRDKRDARRELKLLILGTGSGKSTFIKQRIIHHGAGYSE 63
 QY 73 EERKGRPLVYQNIYFMSRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 130
 Db 64 EDRKGRPLVYQNIYFMSRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 123
 QY 131 YAAAMQWLWDAGIRACVYERRREFHLLDSAVYVLSHLERITEEGYVPTAQDVLRSRMPPT 190

```

124 VSYIKELWESGQECVDRREYQLTDSAKYILSDLRLAVDPYLPTEQDILRVPTT 183
191 GINCYFVQKTNLRIVDVGGQSKERKWIHCFFENFVIALIYSLSEYDQCLENNQNR 250
184 GIIEYFDFLEQIIFRVDVGGQSRERRKWIHCFFENVTIMELVALSEYDQVLVECDNR 243
251 MKESALFGTILELPWFKSTSVILFLNKTIDLEKIPSHLATYFSPFGQPKQDAEAKR 310
244 MEESKALFRTITYPWFNTSSVILFLNKKDLLEKILYSLHLDYFPEYDGGPRDPIARE 303
311 FILDYTRMYTCVDPGSGKARSRLFSHYTCATDQNIKRVKFDVRSVLARYLDE 370
304 FLKMFVDL-----NPDADK-----IYSHFTCATDQNIKRVFAAVKOTILQHNLKE 351
371 INLL 374
352 YNLV 355

RESULT 8
ID Q9X2V4 PRELIMINARY; PRT; 355 AA.
AC Q9X2V4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE GQ PROTEIN, ALPHA SUBUNIT.
GN GQ.
OS Geodia cydonium (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Astrophorida; Geodiidae; Geodia.
OX NCBI_TaxID=6047;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98119577; PubMed=9459489;
RA Seack J., Kruse M., Mueller W.E.G.;
RT "Evolutionary analysis of G-proteins in early metazoans: cloning of
alpha- and beta-subunits from the sponge Geodia cydonium.";
RL Biochim. Biophys. Acta 1401:93-103(1998).
DR EMBL; Y14248; CAB43527.1; -.
DR HSSP; P04896; LAZT.
DR InterPro; IPR001019; -.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
SQ SEQUENCE 355 AA; 41363 MW; 6828C29643F2CE91 CRC64;

```

```

Query Match 44.7%; Score 876.5; DB 5; Length 355;
Best Local Similarity 47.1%; Pred. No. 1.1e-64;
Matches 169; Conservative 73; Mismatches 104; Indels 13; Gaps 2;

Qy 14 LTEDKAAARVDQEIKNRILKQKQDGEUKLLLLGPGSGKSTFTKQRIHGCYSEE 73
Db 5 LSEERLQKRINRINRELQDQKKEIKLLLTGSGSKSTFTKQRIHGCYSG 64
Qy 74 ERKCFRLVYQNFVSMRAMIEMERLQIPFSPESKHHASLVMSODPYKVTTFEKRYAA 133
Db 65 DCLEYKNLVFNILMSHMSLQTAELKXAYIDPDQRHVQLLMALRPETAQSLGGTC 124
-Qy 134 AMQWLWDAGIRACYERRERFHLDSAVVYLSHLERITEGYPTAODVLRSMPTTGIN 193
Db 125 AIRKLWDAGVQCYQRNREYQLSDSKYKYLDDLPRISSNDYVPTTQDVLVRVPTTGIN 184
Qy 194 EYCSVQKTNLRIVDVGGQSKERKWIHCFFENFVIALIYSLSEYDQCL-EENNQENRMK 252
Db 185 EYPTINKIIFKMWVGGQSRERRKWIHCFFENVTIMELVALSEYDQILVEADSVNRMV 244
Qy 253 ESLALFGTILELPWFKSTSVILFLNKTIDLEKIPSHLATYFSPFGQPKQDAEAKRFI 312
Db 245 ESLHLFNTIISYPWFKSSITLLFNKKDLLEKVMHSHLDIYFEEDYDGPCKDHVSARESI 304

```

```

Qy 313 LDMYTRMYTCVDPGSGKARSRLFSHYTCATDQNIKRVKFDVRSVLARYLDEI 371
Db 305 AKMFI-----SINDMRSADIYPHFTCATDQNIKRVFVYVKNHILQOQHTEV 351

RESULT 9
ID Q9Y207 PRELIMINARY; PRT; 305 AA.
AC Q9Y207;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE G PROTEIN A SUBUNIT 3 (FRAGMENT).
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99246375; PubMed=10229568;
RA Suga H., Koyanagi M., Hoshiyama D., Ono K., Iwabe N., Kuma K.,
RA Miyata T.;
RT "Extensive gene duplication in the early evolution of animals before
the parazoan-eumetazoan split demonstrated by G proteins and protein
tyrosine kinases from sponge and hydra.";
RL J. Mol. Evol. 48:646-653(1999).
DR EMBL; AB006541; BAA81695.1; -.
DR HSSP; P10824; IBOF.
DR InterPro; IPR001019; -.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
DR NON_TER 1
FT SEQUENCE 305 AA; 35491 MW; 2639884D438586E9 CRC64;
SQ SEQUENCE 305 AA; 35491 MW; 2639884D438586E9 CRC64;

```

```

Query Match 42.3%; Score 829.5; DB 5; Length 305;
Best Local Similarity 52.1%; Pred. No. 7.1e-61;
Matches 165; Conservative 57; Mismatches 78; Indels 17; Gaps 4;

Qy 63 RIHHCAGYSEBERKGFRLVYQNFVSMRAMIEMERLQIPFSPESKHHASLV-----MS 118
Db 1 RIHHCQGVTEDDRGGYTNLVFLNIYQAOALTRAMNLKISYSPANEENARLILDVDSL 60
Qy 119 QDPYKVT-TFEKRYAAAMQWLWDAGIRACYERRERFHLDSAVVYLSHLERITEGYVP 177
Db 61 EDKSRITITLSPQVASAIESLWKDSGQIEVDYDRREYQLSDSAKYLSDLKRICAPNVP 120
Qy 178 TAQDVLRSRMPPTTGINEYCFVQKTNLRIVDVGGQSKERKWIHCFFENFVIALIYSLSE 237
Db 121 TMQDVLRAAPTGTGIIIEYFDFLDTIIFRMVDVGGQSRERRKWIHCFFENVTIMELVALSE 180
Qy 238 YDQCLEENNQENRMKESLALFQITILELPWFKSTSVILFLNKTIDLEKIPSHLATYTPPS 297
Db 181 YDQVLFESQENRMDESKALFKTITITYPWFLLQSSITLLFNKTIDLEKIQKSDLTQYPE 240
Qy 298 FQGPQDAEAAKRFILDMYTRMYTCVDPGSGKARSRLFSHYTCATDQNIKRVKFK 357
Db 241 YDGAKGDAKAAAEFLKMFVDL-----NPDYDK-----IYSHFTCATDQNIKRVFA 288
Qy 358 DVYRDSVLARYLDEINLL 374
Db 289 AVRDITLQNLKAYNLV 305

RESULT 10
ID Q9I7C8 PRELIMINARY; PRT; 279 AA.
AC Q9I7C8;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE G PROTEIN ALPHA 49B.

```

G-ALPHA-49B OR CGL17759.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 *RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch L., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Rosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.C.,
 RA Fairclat R., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
 RA Shue B.C., Sidenklovas I., Simpson M.C., Skupski M.P., Shen T.,
 RA Spraker E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Strizkas R., Turner C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang J., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Willyam S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Zeng R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheg X.H., Zhong E.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome of Drosophila melanogaster";
 RL Science 287:2185-2195 (2000).
 DR EMBL: AF003825; GenBank: AF003825; J: -;
 DR InterPro: IP0001013; : 2;
 DR Pfam: PF00503; G-alpha: 2;
 DR PRINTS: PR00319; GPROTEINA.
 DR SEQUENCE 279 AA; 32844 MW; A9609C515FF24796 CRC64;

Query Match 40.3%; Score 790; DB 5; Length 279;
 Best Local Similarity 55.1%; Pred. No. 1.2e-57;
 Matches 151; Conservative 51; Mismatches 54; Indels 18; Gaps 2;

QY 13 CLTDEKAARVDQETNRIILLKQKQDRELKLLILGPGESGKSTFIKQRIIHGAGYSE 72
 DB 13 CLTDEKAARVDQETNRIILLKQKQDRELKLLILGPGESGKSTFIKQRIIHGAGYSE 72
 QY 73 EERKGFRLVYQNIQVSPAMTAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 132
 DB 73 EERKGFRLVYQNIQVSPAMTAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 132
 QY 64 EDRGYKILVFQNIQVSPAMTAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 105
 DB 64 EDRGYKILVFQNIQVSPAMTAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 105
 QY 133 AAMWLWRDAGTRACYERREFFHLLDSAVYLLSHLERITEEGYVPTAQDVLRSRPTTGI 192
 DB 133 AAMWLWRDAGTRACYERREFFHLLDSAVYLLSHLERITEEGYVPTAQDVLRSRPTTGI 192
 QY 106 NAIKTLWDAGTQECYDRRREYQVLTDSAKYLLKDRVAQPAFLPTEQDILRVRPPTGI 165
 DB 106 NAIKTLWDAGTQECYDRRREYQVLTDSAKYLLKDRVAQPAFLPTEQDILRVRPPTGI 165
 QY 193 NEYCFSVOKTNLRIVDVGQSKERKWKTHCFENVIALIYLSLSEYDQCLENNQENRMK 252
 DB 193 NEYCFSVOKTNLRIVDVGQSKERKWKTHCFENVIALIYLSLSEYDQCLENNQENRMK 252

DB 166 IEYFPDLEEIRFMVMDVGQSRERKWKTHCFENVIALIYLSLSEYDQILFESDNENRLE 225
 QY 253 ESLALFGTILELPWFKSTSVILFLNKTDLLEIKI 286
 DB 226 ESKALFTIITFEWFKNKSIIILFNKMDVLEIKI 259
 NCBI_TaxID=7227;
 RESULT 11
 Q9HF99 PRELIMINARY; PRT; 354 AA.
 AC Q9HF99;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE ETEROTRIMERIC G PROTEIN ALPHA SUBUNIT B.
 GN SCOP-B.
 OS Schizophyllum commune (Bracket fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Stereales;
 OC Schizophyllum commune; Schizophyllum.
 OX NCBI_TaxID=5934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamagishi K., Kimura T., Suzuki M., Shimoto H.;
 RT "Schizophyllum commune heterotrimeric G protein alpha subunit B";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB051303; BAB18736; : -;
 DR SEQUENCE 354 AA; 40448 MW; 4BC2E5DF0CF36654 CRC64;

Query Match 40.2%; Score 787.5; DB 3; Length 354;
 Best Local Similarity 44.4%; Pred. No. 2.6e-57;
 Matches 160; Conservative 71; Mismatches 112; Indels 17; Gaps 6;

QY 17 DEKAAARVDQETNRIILLKQKQDRELKLLILGPGESGKSTFIKQRIIHGAGYSEERK 76
 DB 17 DEKAAARVDQETNRIILLKQKQDRELKLLILGPGESGKSTFIKQRIIHGAGYSEERK 76
 QY 10 DEEAKARNDEIESQLAKKDRMAKNEIKMLLJAGESGKSTVLQKMLIHGGTNEQERD 68
 DB 10 DEEAKARNDEIESQLAKKDRMAKNEIKMLLJAGESGKSTVLQKMLIHGGTNEQERD 68
 QY 77 GERPLVYQNIQVSPAMTAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYAAA 134
 DB 77 GERPLVYQNIQVSPAMTAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYAAA 134
 QY 69 SYKEITFNTIQSMRAILEAMPQDILPT-PQDARRAVINSM-PMQIADVLPPDVDA 126
 DB 69 SYKEITFNTIQSMRAILEAMPQDILPT-PQDARRAVINSM-PMQIADVLPPDVDA 126
 QY 135 MOWLWRDAGTRACYERREFFHLLDSAVYLLSHLERITEEGYVPTAQDVLRSRPTTGI 194
 DB 135 MOWLWRDAGTRACYERREFFHLLDSAVYLLSHLERITEEGYVPTAQDVLRSRPTTGI 194
 QY 127 IRSLLWDPGVQRAQVRSRQFQNDLSAVYFENSIDRAAPNPMPTDQILRSRYKTTGTE 186
 DB 127 IRSLLWDPGVQRAQVRSRQFQNDLSAVYFENSIDRAAPNPMPTDQILRSRYKTTGTE 186
 QY 195 YCFSVOKTNLRIVDVGQSKERKWKTHCFENVIALIYLSLSEYDQCLENNQENRMK 254
 DB 195 YCFSVOKTNLRIVDVGQSKERKWKTHCFENVIALIYLSLSEYDQCLENNQENRMK 254
 QY 187 TTEKVGELTYKLFQDVGQSRERKWKTHCFENVIALIYLSLSEYDQCLENNQENRMK 246
 DB 187 TTEKVGELTYKLFQDVGQSRERKWKTHCFENVIALIYLSLSEYDQCLENNQENRMK 246
 QY 255 LALFGTILELPWFKSTSVILFLNKTDLLEIKI-TSHLATYFSPFGPKODAAKRFILD 314
 DB 255 LALFGTILELPWFKSTSVILFLNKTDLLEIKI-TSHLATYFSPFGPKODAAKRFILD 314
 QY 247 LTLFSDICNSRWFVKTSIIILFNKIDLEFAEKL-RSPGLGDFPDYTG-GDNYDAACDYLLH 305
 DB 247 LTLFSDICNSRWFVKTSIIILFNKIDLEFAEKL-RSPGLGDFPDYTG-GDNYDAACDYLLH 305
 QY 315 MYTRMYTGVGDPGEGSKGARSRRFLSFHYTCA-DTONIRKVFQDVRDVLARVLDLEILL 374
 DB 315 MYTRMYTGVGDPGEGSKGARSRRFLSFHYTCA-DTONIRKVFQDVRDVLARVLDLEILL 374
 QY 306 REVSL-----NQSAATQIVAHYTCA-DTQOIKFVLSAIDILLQLHLRECGLL 354
 DB 306 REVSL-----NQSAATQIVAHYTCA-DTQOIKFVLSAIDILLQLHLRECGLL 354

RESULT 12
 Q9HFA2 PRELIMINARY; PRT; 353 AA.
 AC Q9HFA2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE G PROTEIN ALPHA SUBUNIT.
 GN WGA2.
 OS Rosellinia necatrix.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Xylariales; Xylariaceae; Rosellinia.
 OX NCBI_TaxID=77044;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-W8;

RA Aimi T., Sanae K., Wang Q., Morinaga T.;
 RT "Molecular cloning of three genes for G protein alpha-subunit-protein
 from white root rot fungus, *Rosellinia necatorix*.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB045578; BAB20820.1; -;
 SQ SEQUENCE 353 AA; 41016 MW; 0EB84C4B292B70E6 CRC64;

Query Match 39.0%; Score 764.5; DB 3; Length 353;
 Best Local Similarity 42.7%; Pred. No. 2.1e-55;
 Matches 152; Conservative 67; Mismatches 120; Indels 17; Gaps 5;

QY 15 TEDEKAARVDQENRILLQKQKDRGELKLLLLGPGESGKSTFIKQMRHHGAGYSEE 74
 DB 7 TEEKEGAR-NEETENQKRDQKMMORNEIKMLLGGESGKSTILKQMKLHGGYSRDE 65
 QY 75 RKGFRLVYQNIQFVSMRAMIEMERLQIPFSRPSKHHASIVMSQDPYK--TTFEKRYA 132
 DB 66 RESFKEIFNTVQSMRVILEAMESLEPLPEPRMEYHVQTIEMQ-PAQIEGDVLPPEVG 124
 QY 133 AAMQWLWDAGIRACYYRRERREHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPPTGI 192
 DB 125 GALEALWKDHGVQCFKRSREYQLNDSARYFDNIVRIATPDYMPDQDVLRSRVKTGTI 184
 QY 193 NEYCFVQKTNLRIVDVGQKSERKKWHCFENVIALIYLASLSEYDQCLLENQENRMK 252
 DB 185 TETTFIIGDLTYRMDFVGGQSRERKKWHCFENVITILFLVAISEYDQCLLENQENRMK 244
 QY 253 ESALFQGTILELPWFKSTSVILFLNKTDLIEEKIPTSHLATYFPSPGPKQDAEAAKRFI 312
 DB 245 EALTLDISICNSRWFIKTSILFLNKLIDRFKEKLPVSPMKNYFPDYE-GDDYAAACDYI 303
 QY 313 LDMYTRMYTCVDPGEGSKGARSRRFLSHYTCATDTQNIKRVKDFVDSVLARYL 368
 DB 304 LNRFV-----SLNQHETKQIYTHETCATDTTQIRFVMAAVNDIIQENL 347

RESULT 13

Q9NLP9 ID Q9HP9 PRELIMINARY; PRT; 353 AA.
 AC Q9HP9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE G PROTEIN ALPHA SUBUNIT.
 GN G1.
 OS Botrytis cinerea (Botryotinia fuckeliana).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
 OC Helotiales; Sclerotiniaceae; Botryotinia.
 OX NCBI_TaxID=40559;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SAS56;
 RA Kasulke D., Tudzynski P., Tudzynski B.;
 RT "Cloning and characterisation of genes coding for G-alpha subunits
 from Botrytis cinerea.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y18436; CAC19871.1; -;
 SQ SEQUENCE 353 AA; 40983 MW; 01746CEID25ABFAF CRC64;

Query Match 38.5%; Score 755.5; DB 3; Length 353;
 Best Local Similarity 42.4%; Pred. No. 1.2e-54;
 Matches 151; Conservative 69; Mismatches 119; Indels 17; Gaps 5;

QY 15 TEDEKAARVDQENRILLQKQKDRGELKLLLLGPGESGKSTFIKQMRHHGAGYSEE 74
 DB 7 TEEKEGAR-NEETENQKRDQKMMORNEIKMLLGGESGKSTILKQMKLHGGYSRDE 65
 QY 75 RKGFRLVYQNIQFVSMRAMIEMERLQIPFSRPSKHHASIVMSQDPYK--TTFEKRYA 132
 DB 66 RESFKEIFNTVQSMRVILEAMESLEPLDDQRAEYHVQTIEMQ-POQIEGDVLPPEVG 124

QY 133 AAMQWLWDAGIRACYYRRERREHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPPTGI 192
 DB 125 SAIAALWKDHGVQCFKRSREYQLNDSARYFDNIERIAQHDYMPDQDVLRSRVKTGTI 184
 QY 193 NEYCFVQKTNLRIVDVGQKSERKKWHCFENVIALIYLASLSEYDQCLLENQENRMK 252
 DB 185 TETTFIIGDLTYRMDFVGGQSRERKKWHCFENVITILFLVAISEYDQCLLENQENRMK 244
 QY 253 ESALFQGTILELPWFKSTSVILFLNKTDLIEEKIPTSHLATYFPSPGPKQDAEAAKRFI 312
 DB 245 EALTLDISICNSRWFIKTSILFLNKLIDRFKEKLPVSPMKNYFPDYE-GDDYAAACDYI 303
 QY 313 LDMYTRMYTCVDPGEGSKGARSRRFLSHYTCATDTQNIKRVKDFVDSVLARYL 368
 DB 304 LNRFV-----SLNQHETKQIYTHETCATDTTQIRFVMAAVNDIIQENL 347

Query Match 38.5%; Score 754; DB 5; Length 357;
 Best Local Similarity 43.3%; Pred. No. 1.6e-54; Indels 30; Gaps 6;

QY 9 CCPWCLTEDEKAARVDQENRILLQKQKDRGELKLLLLGPGESGKSTFIKQMRHHG 68
 DB 3 CAP-SKSENDKDAVSKSKEIDKQKDAENARKEVKLLLLGAGESGKSTIAQMKILHQD 61
 QY 69 GYSEERKGRPLVYQNIQFVSMRAMIEMERLQIPFSRPSKHHASIVMSQDPYK---- 124
 DB 62 GFSEERKGNFAPVYVNTTQISWVAIVKAMESLIGDIYEDNERE-----EDGRKIRAS 113
 QY 125 TTFEK-----RYAAMQWLWDAGIRACYYRRERREHLLDSAVYLSHLERITEEGYVPTA 179
 DB 114 QTMEEIDITADIGALKRLMDKGVACVARSREYQLNDSARYLQDLCSPDIIPTE 173
 QY 180 QDVLRSRMPPTGTINRYCFVQKTNLRIVDVGQKSERKKWHCFENVIALIYLASLSEYD 239
 DB 174 QDVLTRVKTGTIETTFQYKKNLFTLIDVGQSRERKKWHCFQDVTAILFCVAMSAYD 233
 QY 240 QCLLENQENRMKESLALFGTILELPWFKSTSVILFLNKTDLIEEKIPTSHLATYFPSPQ 299
 DB 234 QVLADEDETNRMIESLKLFSQICNNPFPAKTSMLFLNKKDLFEKIKKSPINTCFKBEY 293
 QY 300 GPKQDAEAAKRFILDMYTRMYTCVDPGEGSKGARSRRFLSHYTCATDTQNIKRVKDV 359
 DB 294 GENSADASEH-IKEQF-----EAAKNKNCMEKIYTHETCATDTGNRFVDAV 341

[illegible]

Search completed: September 6, 2001, 10:55:04
Job time: 84 sec


```

151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
152 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
451 CGCGGGAATTCACACCTGCTCGATTACGCGGTACTACCTGTCCACCT 500
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
168 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
501 GGAGCGCATCACCGAGAGGGCTACGTCACACAGCTCAGGACGTGCTCC 550
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
185 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
551 GCAGCCGATGCCACCTGCTGCTACACAGTACTGCTTCTCGCTGACG 600
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
202 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
601 AAAACCACTGCGGATCGTGAGAGCTCGGGGCGGAGGAGGAGGAGGTA 650
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
218 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
651 GAAATGATCCATTGTTTCGAGAACGTGATCGGCCCTCATCTACCTGG 700
234 erLeuSerGluTyrAspGlnCysLeuGluAsnGlnAsnGlnGluAsnArg 250
251 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
701 CACTGATGAATACGACCACTGCTGAGAGAGAACAAACAGGAGAACCG 750
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrp 267
268 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
751 ATGAAGGAGAGCTCGCATTTGTTGGGACTATCTCGGAACCTACCTGG 800
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGlu 284
285 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
801 CAAAGCACATCGCTCATCTCTCTTCACAAACAAACGACATCTGAGG 850
284 lLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
285 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
851 AGAAATCCCCACCTCCACCTGCTACCTATTTCCTCCAGTTCCAGGGC 900
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
318 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
901 CCTAAGCAGATGCTGAGGAGGCAAGAGGTTTCATCTGGACATGTACAC 950
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGly 334
335 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
951 GAGATGTACACCGGGTGTGGAGCGCCCGGAGGCGGAGCAAGAGGGCG 1000
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
351 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1001 CAGATCCCGACGCTTTTCAGGCACCTACACATCTGCCACAGACACAG 1050
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
368 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1051 AACATCCGCAAGGTCTTCAAGGAGCTGCGGGAGCTCGGTGCTCGCCGCT 1100
367 rLeuAspGluIleAsnLeuLeu 374
375 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1101 CCTGGAGAGATCAACCTGCTG 1122

```

seq_name: gb_pat1:AR106753

```

seq_documentation_block:
LOCUS AR106753 2060 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6107091.
ACCESSION AR106753
VERSION AR106753.1 GI:12821283
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2060)
AUTHORS Cowbert, L.M.
TITLE Antisense inhibition of G-alpha-16 expression
JOURNAL Patent: US 6107091-A 1 22-AUG-2000;
FEATURES
Location/Qualifiers

```

```

source 1..2060
BASE COUNT 415 a 631 c 640 g 374 t
ORIGIN

alignment_scores:
Quality: 1960.00 length: 374
Ratio: 5.241 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-878-801-2 x AR106753
Align seg 1/1 to: AR106753 from: 1 to: 2060

1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
220 ATGGCCCGCTCGCTGACCTGGCGCTGCTGCCCTGGGTGCCCTGACGGAGGA 269
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeu 34
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
270 TGAGAGGGCGCGCGCGGGTGGACACGAGATCAACAGGATCCTCTTGG 319
34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
320 AGCAGAAGAAGACGAGACCGCGGGGAGCTGAAGTCGCTGCTTTGGGCCCA 369
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
370 GCGGAGAGCGGGAAGACACCTTCATCAAGCAATGCGGATCATCCACGG 419
67 yAlaGlyTyrSerGluGluArgLysGlyPheArgProLeuValTyrG 84
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
420 CGCGGCTACTCGAGGAGGAGCGCAGGGCTCCGGCCCTGGTCTACC 469
84 lAsnIlePheValSerMetArgAlaMetIleIleAlaMetGluArgLeu 100
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
470 AGAACATCTTCGTGCTCCATCGGGCCATGATCAGGCCATGGAGCGCTG 519
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
520 CAGATTCATTCAGACGCGCGCGAGAGCAAGCAACACGCTAGCCTGGTCA 569
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
570 GAGCCAGGACCCCTATAAAGTGACACGCTTGAGAAGCGCTACGCTCGG 619
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
620 CCATGCACTGGCTGTGGAGGATCGCGGCATCAGGGCTGCTATGAGCGT 669
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
670 CGCGGGGAATTCACCTGCTCGATTACGCGCTTTACTACCTGTCCACCT 719
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
720 GGAGCGCATCACCGAGGAGGCTACGTCGCCCAAGCTCAGGACGTGCTCC 769
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
770 GCAGCCGATGCCACCTGCGGATCAACAGGACTACTGCTTCTCCGTCAG 819
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
820 AAAACCAACCTGCGGATCGTGAGAGCTCGGGGCGGAGGAGGAGGTA 869
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
870 GAAATGATCCATTGTTTCGAGAACGTGATCGGCTCCTCATCTACCTGG 919
234 erLeuSerGluTyrAspGlnCysLeuGluAsnGlnGluAsnArg 250

```

```

|||||
920 CACTGAGTGAATACGACCACTGCTGGAGGAGAACCAACGAGGAGACCGC 969
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTyrPhe 267
|||||
970 ATGAAGAGAGCCCTCCGATGTTTGGGACTATCTGGAACCTACCTGGTT 1019
|||||
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
1020 CAAAAGCACATCGCTCATCTCTTCTCAACAAACCGACATCTCTGGAGG 1069
|||||
284 LuLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
1070 AGAAATCCCACTCCCTCCCTGGCTACCTATTTCCCACTTCCAGGGC 1119
|||||
301 ProLysGlnAspAlaGluAlaAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
1120 CCTAAGCAGGATGCTGAGGAGCAAGAGGTTTCATCTGGACATGTACAC 1169
|||||
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
|||||
1170 GAGGATGTACACCGGFGCTGACGCGCCGAGGAGGAGCAAGAGGGCG 1219
|||||
334 laArgSerArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1220 CACGATCCCGACGCTTTTCCAGCCACTACACATGTGCCACACACACAG 1269
|||||
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
|||||
1270 AACATCGCAAGGCTTCAAGAGAGCTGCGGAGCTCGGTGCTGCCCGCTA 1319
|||||
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1320 CCTGGACGAGATCAACCTGCTG 1341

seq_name: gb_pr10:HUMGAL6

seq documentation block:
LOCUS HUMGAL6 2060 bp mRNA PRI 27-APR-1993
DEFINITION Human G-alpha 16 protein mRNA, complete cds.
ACCESSION M63904
VERSION M63904.1 GI:182891
KEYWORDS G-alpha 16 protein.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Anatrua,T.I.II., Steele,D.A., Slepak,V.Z. and Simon,M.I.
G-alpha16, a G protein alpha subunit specifically expressed in
hematopoietic cells
Proc. Natl. Acad. Sci. U.S.A. 88, 5587-5591 (1991)
JOURNAL 91288509 Location/Qualifiers
MEDLINE
FEATURES
source
1..2060
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HL-60"
<1..2060
/product="G-alpha-16 protein"
220..1344
/codon_start=1
/product="G-alpha-16 protein"
/protein_id="AAA35860.1"
/db_xref="GI:182892"
/translation="MARSLWRCPWCLTEDEKAAARVDOEINRILLEOKKODRGELK
LLLGPGESKSTFIKQMLIHGAGYSEERKGRFLVIGNIFVSMRAMEARLQI
PFSRPSKHSALVMQDPYKVTTFKKRYAAAMWLWRDAGRACRYERRRPFLLDQA
VYLSHNERITEGYVPTAQLVLRSMPTTGINCYCFVQKTNLRIDVDGQKSERKK
WIHFENVALIVLASLSEYDOCLENNQENRMKESLALFTILELPWFKSTSVILFL
NKTDILEEKIPTSLATYFPFSGPKODAEAAKRFILDMYTRMTGCDVDPESKKGA
RSRLFSHYTCARDTQNIKRVKFDVDRSVLARIYDEINLL"
BASE COUNT 415 a 631 c 640 g 374 t

```

```

ORIGIN
alignment_scores:
  Quality: 1960.00 Length: 374
  Ratio: 5.241 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
  US-08-878-801-2 x HUMGAL6
Align seg 1/1 to: HUMGAL6 from: 1 to: 2060
1 MetAlaArgSerLeuThrTyrPheArgCysCysProTyrPheCysLeuThrGluAs 17
220 ATGGCCCGCTCGCTGACCTGGCGCTGCTGCCCTGGTGCCTGACGAGGAGGA 269
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeuG 34
270 TCAGAAGCGCCGCCCGGGTGGACGAGATCAACAGGATCCTCTTGG 319
34 LuGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
320 ACAGAAGACGAGACCGCGGGGAGCTCAAGCTGCTGCTTTGGGCCCA 369
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
370 GCGAGAGCGGGAAGACACCTTCATCAAGCAGATCGGATCATCCACGG 419
67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
420 CGCGGGCTACTCGGAGGAGGAGCGCAAGGGCTTCGGGCCCTGGTCTACC 469
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
470 AGAACATCTTCGTGTCATCGCGGCATGATCGAGGCCATGGAGCGGTG 519
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
520 CAGATTTCATTCAGCAGCGCCGAGCAAGCAGCACCGCTAGGCTGGTCA 569
117 tSerGlnAspProTyrLysValThrThrPheGluLysArgTyrAlaAla 134
570 GAGCCAGGACCCCTATAAAGTGACACAGCTTGAGAAAGCGCTACGCTGCGG 619
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
620 CCATGCAGTGGCTGTGGAGGATCGCGGCATCGCGGCCTGCTATGAGCGT 669
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
670 CGCGGGGAATTCACCTGCTCGATTGAGCCGCTGCTACTACTCTGCCACCT 719
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
720 GGAGCGCATCCCGAGGAGGGCTACGTCCCCACAGCTCAGGACGTGCTCC 769
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
770 GCAGCGCATGCCACCACTGGCATCAACGAGTACTGCTTCTCCGTCAG 819
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
820 AAAACCAACCTGCGGATCGGTGGACGCTCGGGGGCCAGAGTCAGAGCGTAA 869
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
870 GAATGGATCCATGTTTTCAGAAACGATGATCGCCCTCATCTACCTGSCCT 919
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnAsnGlnGluAsnArg 250
920 CACTGAGTGAATACGACCACTGCTGGAGGAGAGAACACAGGAGAACCCG 969

```

251 MetLeuSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
 |||
 970 ATGAAGAGAGCTCCATGTTGGACTATCTTGGAACTACCCGGTT 1019
 267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
 |||
 1020 CAAAGACATCCGTCATCTCTTCTCAACAAACCCGACATCCTGGAGG 1069
 284 LuLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
 |||
 1070 AGAAATCCCACTCCACCTCCGCTGCTACTATTTCCCAAGTTTCCAGGGC 1119
 301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
 |||
 1120 CTAAGCAGGATGCTGAGCAGCAAGAGTTTCATCTGGACATGATAC 1169
 317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysGlyA 334
 |||
 1170 GAGGATGATACCCGGTGGTGGAGCCCGCCGAGGAGCAAGAGGGG 1219
 334 LaArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
 |||
 1220 CACGATCCGAGCGCTTTTACGCCACTACATGTCACACACACACAG 1269
 351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
 |||
 1270 AACATCCGCAAGTCTTCAAGGACGTGGGAGCTCGGTGCTCGCCGCTA 1319
 367 rIeuAspGluIleAsnLeuLeu 374
 |||
 1320 CTGGACGAGATCAACCTGCTG 1341

seq_name: gb_rol:AB015308

seq_documentation_block:
 LOCUS AB015308 1766 bp mRNA ROD 06-FEB-1999
 DEFINITION Rattus norvegicus mRNA for GTP binding protein alpha 15, complete cds.

ACCESSION AB015308
 VERSION AB015308.1 GI:3219264
 KEYWORDS GTP binding protein alpha 15.
 SOURCE Rattus norvegicus tissue_lib:tongue epithelium cDNA to mRNA.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1766)
 AUTHORS Kusakabe,Y., Yamaguchi,E., Tanemura,K., Kameyama,K., Chiba,N., Arai,S., Emori,Y. and Abe,K.
 TITLE Direct Submission
 JOURNAL Submitted (05-JUN-1998) to the DDBJ/EMBL/GenBank databases. Yoko Kusakabe, University of Tokyo Dept. Applied Biol. Chem.; 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan (E-mail:aka7308@hongo.ecc.u-tokyo.ac.jp, Tel:81-3-5802-8897, Fax:81-3-5802-8897)

REFERENCE 2 (sites)
 AUTHORS Kusakabe,Y., Yamaguchi,E., Tanemura,K., Kameyama,K., Chiba,N., Arai,S., Emori,Y. and Abe,K.
 TITLE Identification of two alpha-subunit species of GTP-binding proteins, Galphas and Galphag, expressed in rat taste buds
 JOURNAL Biochim. Biophys. Acta 1403 (3), 265-272 (1998)
 MEDLINE 98351893
 FEATURES Location/Qualifiers

1..1766
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /tissue_lib="tongue epithelium"
 231..1355
 /codon_start=1
 /product="GTP binding protein alpha 15"
 /protein_id="BAA28827.1"
 /db_xref="GI:3219265"
 /translation="MARSLTWCCPWCILTEEFKTAARDIOEINKLLLEOKQBERGELK
 LLLPGESGSKSTFIKQMRIIHGAGYSEEDRFRFLIYYQNFVSMQAWIEAMDRLQI

PFSPDQKSHASLVMTQDPY<VSSFEKPYAVAMOYLWRDAGIRACIYERRRFFHLLDSA
 VYLSHLIERAEDDYIPTAQDVLRSRMPTTGINNEYCFSVOKTKLRIVDAGGOKSERKK
 WIFCFENVIALIYLASLSEY)DCLSENSQENRMKESLAFSTILELPWFKSTSVILFL
 NKTDILEDKIHTSLSAYFP:FOGPRRDAAEAARFILDIMYARVYASCAEPHDGGRKGS
 RARLFAHETCATDTHSVRS/FKDVDRDVLARYLDEINLL"
 1766

polyA_site

BASE COUNT 401 a 516 c 513 g 336 t
 ORIGIN

alignment_scores:
 Quality: 1722.00 length: 374
 Ratio: 4.692 gaps: 0
 Percent Similarity: 98.128 Percent Identity: 86.096

alignment_block:
 US-08-878-801-2 x AB015308 ..

Align seg 1/1 to: AB015308 from: 1 to: 1766

1 MetAlaArgSerLeuThrTyrPheArgCysCysProTyrCysLeuThrGluAs 17
 |||
 231 ATGGCCCGTCCCTGACTTGGGGCTGCTGCTCCCTGGTGCCTCACGAAGA 280
 17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeuG 34
 |||
 281 GGAGAAGACTGCGCCAGAAATCGACCAGGAGA1CAACAAGATTTTTGTGG 330
 34 LuGlnLysLysGlnAspArgGlyGluLeuLysIleLeuLeuLeuGlyPro 50
 |||
 331 AACAGAAGAAACAGAGCGCGGGGAATTGAAAC TCCTGCTGTGGGGCCC 380
 51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
 |||
 381 GGTGAGACGGGAAAGACACGTTTCATCAAGCAPATGCCATCATTCACGG 430
 67 ValAcgTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
 |||
 431 CGCCGGCTACTCTGAGGAGGACCGAGAGCCTTCGGCTGCTGCTACCC 480
 84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
 |||
 481 AGAACATCTTCTGCTCCATGAGGCCCATGATTGAAGCAATGCAGAGGCTG 530
 101 GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValme 117
 |||
 531 CAGATCCCTTCAGACGGCCGAGCAAGAAACAGACCGCCAGCTGGTGAT 580
 117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
 |||
 581 GACCCAGAGCCCTTATAAGTGAGCTGCTTCGA3AAGCCATATGCAGTGG 630
 134 LaMetGlnTyrLeuTyrPheArgAspAlaGlyIleArgAlaCysTyrGluArg 150
 |||
 631 CCATGCAGTACCTGTGCGCGGACGCGGGCATCCACGATGCTACGACGG 680
 151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
 |||
 681 AGCGGTGAATTCACCTGCTGGACTCCCGGTGACTACCTGTGCACACCT 730
 167 uGluArgIleThrGluGluGlyTyrValProTh::AlaGlnAspValLeu 184
 |||
 731 GGAGCGCATCCGAGGAGCAGCACTACATCCCCAC:GGCAGGAGGCTGCTG 780
 184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
 |||
 781 GCAGTCGATGCCACACCTGGCATCAATGAGTACTGCTTTTCCGTCAG 830
 201 LysThrAsnLeuArgIleValAspValGlyGlyLysSerGluArgly 217
 |||
 831 AAAACCAAACTGCGTCATCGTGGATGCTGGCGGCAGAAAGTCAGAACGTAA 880


```

|||||
601 AAGACCAACCTCGCATCGTGGACGTCGGGGCCAGAAATCGAGCGCAG 650
|||||
217 slystrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaAs 234
|||||
651 GAAGTGGATGCCACTGCTTCGAGGAGCTGACGGCGCTCATCTTCCTGGCCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnAsnGlnGluLeuAsnArg 250
|||||
701 CACTGAGCGAGTACGACAGTGCCTGGAGGAGACGGCCAGGAGAACCGC 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrrpPh 267
|||||
751 ATGCAGGAGAGCCTGGCGCTGCTCGGCACCGCTCGTGGCGCTTCGCCCTGGTT 800
|||||
267 elysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CCGGCCACCTCGGTATCTCTCTTCCTCACAGACGACATCCCTGGAGG 850
|||||
284 luLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
851 ACAAGGTCGCGACGTCGCGACCTGCTCCGCGGCTTCGCGGGG 900
|||||
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
901 CCCCCAGGACCGCGAGCGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 950
|||||
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
|||||
951 ACGCGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1000
|||||
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1001 CGCGCTCCCGCGCGCTTCAGGACCTACACGTGCGCGCACGACGACGACG 1050
|||||
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
|||||
1051 AACATACGCAAGGTCTCAAGGACGTCGCGGACTCGGTGCTCGCGCGCTA 1100
|||||
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1101 CCTGCAGCAGATCAACCTGCTG 1122

```

seq_name: gb_pat1:AR095767

```

seq_documentation_block: 1125 bp DNA PAT 08-SEP-2000
LOCUS AR095767
DEFINITION Sequence 3 from patent US 6004808.
ACCESSION AR095767
VERSION AR095767.1 GI:10023948
KEYWORDS
SOURCE unknown.
ORGANISM unknown.

```

```

REFERENCE
  1 (bases 1 to 1125)
  Negulescu, P., Offermanns, S., Simon, M. and Zuker, C.
  TITLE Promiscuous G-protein compositions and their use
  JOURNAL Patent: US 6004808-A 3 21-DEC-1999;
  FEATURES
    Location/Qualifiers
      1..1125
        /organism="unknown"

```

```

BASE COUNT 263 a 328 c 325 g 209 t
ORIGIN

```

```

alignment_scores:
  Quality: 1705.00
  Ratio: 4.658
  Percent Similarity: 97.861
  Percent Identity: 84.759
alignment_block:
  US-08-878-801-2 x AR095767

```

Align seg 1/1 to: AR095767 from: 1 to: 1125

```

1 MetAlaArgSerLeuThrTrpArgCysProTrrpCysLeuThrGluAs 17
|||||
1 ATGGCCCGGTCCTGACCTGGGGCTGCTGTCCCTGGTGGCTGCACAGAGGA 50
|||||
17 pCulLysAlaAlaAlaArgValAspGlnGluT...eAsnArgIleLeuLeuG 34
|||||
51 GGAGAGAGACTGCGCCAGAAATCGACAGGAGAA:CAACAGGATTTGTTGG 100
|||||
34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuLeuPro 50
|||||
101 AACAGAAAAACAAGAGCGCGAGGAATTTGAAATCTCCTGCTGTTGGGGCT 150
|||||
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
|||||
151 GGTGAGAGCGGAGAGTACGTTTCATCAAGCAATGCGCATCATTCACGG 200
|||||
67 yAlaGlyTyrSerGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
201 TGTGGGCTACTCGGAGGAGGACCGCAGAGCCTTCGGGCTGCTCATCTACC 250
|||||
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
251 AGAACATCTTCCTCCATGCGAGCCATGATAGATGCGATGACCGGCTG 300
|||||
101 GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValMe 117
|||||
301 CAGATCCCTTCAGCAGCGCTGCAGCAAGCAACAGCCACGCTAGTGAT 350
|||||
117 tSerGluAspProTyrLysValThrPheGluLysArgTyrAlaAlaA 134
|||||
351 GACCCAGGACCCCTATAAGTGAGCACATTCGGAAGCCATATGACGTGG 400
|||||
134 laMetGlnTrpLeuTrpArgAspAlaGlyIlePrgAlaCystyrGluArg 150
|||||
401 CCATGCACTACCTGTGCGGAGCGCGGCATCCGTGATGCTACGAGCGGA 450
|||||
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 AGCGTGAATTCACCTTCTGGACTCCGCGGTGATTACCTGTCCACCT 500
|||||
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
501 GGAGCGCATATCAGAGGACAGCTACATCCCGACATGCGCAAGACGTGCTGC 550
|||||
184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
551 GCAGTCGATGCGCCACACAGGATCAATGAGTACTGCTTCTCCGTGAAG 600
|||||
201 LysThrAsnLeuArgIleValAspValGlyGly3lnLysSerGluArgLy 217
|||||
601 AAAACCAAACTGCGCATCGTGGATGTTGGTGGCAGAGGTACAGACGTAG 650
|||||
217 sLysTrrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
|||||
651 GAAATGGATTCACCTGTTTCGAGAACGTGATTGCTCTCATCTACTCGGCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
701 CCCTGAGCGAGTATGACCGTGTAGAGGAGATCGATCAGGAGAACCGC 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrrpPh 267
|||||
751 ATGGAGGAGAGTCTCGCTCTGTTTCAGACCATCTAGAGCTGCGCTGGTT 800
|||||
267 elysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CAAGAGCAGCTCGGTCTATCTCTTCTCAACAAACAGGACATCTCTGGAAG 850
|||||
284 luLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
851 ATAAGATTACACCTCCACCTGCGCACATACTCCCGAGCTTCCAGGGA 900

```

[illegible]

seq_name: qb_ro1:MUSGNA15A

seq_documentation_block:					
LOCUS	MUSGNA15A	1353 bp	mRNA	ROD	27-APR-1993
DEFINITION	Mouse G protein alpha subunit (GNA-15) mRNA, complete cds.				
ACCESSION	M80632				
VERSION	M80632.1	GI:193570			
KEYWORDS	G protein alpha-subunit.				
SOURCE	Mus musculus (strain C57BL/6J x DBA F1) (library: lambda ZAP II) adult spleen cDNA to mRNA.				
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. I (bases 1 to 1353)				
REFERENCE	Wilkie,T.M., Scherle,P.A., Strathmann,M.P., Slepak,V.Z. and Simon,M.I.				
AUTHORS	Characterization of G-protein alpha subunits in the Gq class: Expression in murine tissues and in stromal and hematopoietic cell lines				
TITLE					
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053 (1991)				

FEATURES

```

gene
43..1167
/gene="GNA-15"
CDS
43..1167
/gene="GNA-15"
/codon_start=1
/product="G protein alpha s"
/protein_id="AAA37713.1"
/db_xref="GI:193571"

```

[illegible]

```
alignment_scores:
  Quality: 1705.00
  Ratio: 4.658
  Percent Similarity: 97.861
```

```

Length: 374
Gaps: 0
Percent Identity: 84.759

```

```

alignment_block:
  US-08-878-801-2 x MUSGNA15A  ..
  Align seg 1/1 to: MUSGNA15A from: 1 to: 1353

1  MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43  ATGGCCCGGTCCTGACTGTTGGGGCTGCTGCTCCCTGGCTGCACAGAGGA 92
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17  pGluysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeuG 34
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
93  GGAGAAAGACTGCCGCCAGAATCACCAGCAGAGATCAACAGGATTTGTGTG 142
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34  luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuLeuGlyPro 50
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
143  AACAGAAAAACAGAGCGCGAGGAATTGAACATCTCTGTTGTTGGGCGCT 192
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51  GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisGl 67
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
193  GGTGAGAGCGGAGAGTAGCTTCATCAGCAGATGGCATCATTCACGG 242
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67  yAlaGlyTyrSerGluGluuArgLysGlyPheArgProLeuValTyrG 84
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
243  TGTGGGCTACTCGAGGAGGACCGCAGAGACTTCCGGCTGCTCATCTACC 292
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84  InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
293  AGAATCATCTTCGTCTCATGCGAGCCATGATAGATCGATGGACCGGCTG 342
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101  GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
343  CAGATCCCCCTTCAGCAGGCGCTGACACAGCAGCAGCCAGCCAGCTAGT 392
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117  tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
393  GACCCAGGACCCCTATAAAGTGAGCACATTCGAGAGCCCATATGCACTGG 442
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134  laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
443  CCATGCAGTACCTGTGGCGGAGCGCGGCAATCCGTGTCATGCTACGAGCGA 492
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151  ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
493  AGCGGTGAATTCACCTCTTGGAATCCGCGGTGTATTACCTGTGCACACT 542
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167  uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
543  GGAGCGCATATCAGAGACAGCTACATCCCACTGCGCAGACGCTGTCTGC 592
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184  rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
593  GCAGTCGCATGCCACCACAGGCGATCAATGAGTACTGCTCTCCGTGAAG 642
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201  LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
643  AAAACCAACTGCGCATCTGTGGATGTGGTGGCCAGCAGGTCAGAGCGTAG 692
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217  sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
693  GAAATGGATTCACTGTTTCGAGAACGTGATTGCCCTCATCTACCTGGGCT 742
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234  erLeuSerGluTyrAspGlnCysLeuGluGluAsnAsnGlnGluAsnArg 250
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
743  CCTGAGCGAGTAGACCGATGCCCTAGAGGAGAACGATCAGAGAACCGC 792
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251  MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
793  ATGGAGGAGAGTCTCGCTCTCTTACAGCAGCATCTAGAGCTGCCCTGGT 842
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
267  eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
843  CAAGAGCACCTCGGTCACTCTCTTCTTCAACAGCAGCATCTCTTGGAG 892
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



```

284 luLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
893 ATAAAGATTACACACCCACCTGGCCACATATCTCCCGAGCTTCCAGGGA 942
      ::::::::::::::::::::::::::::::::::::::::::::::::::
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
      ::::::::::::::::::::::::::::::::::::::::::::::::::
943 CCCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 992
      ::::::::::::::::::::::::::::::::::::::::::::::::::
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysTysGlyA 334
      ::::::::::::::::::::::::::::::::::::::::::::::::::
993 GCGCGTGTACGCGAGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGCT 1042
      ::::::::::::::::::::::::::::::::::::::::::::::::::
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
      ::::::::::::::::::::::::::::::::::::::::::::::::::
1043 CCGCGCGCGCGCTCTTCGCACACATTCACCTGTGCGCAGCAGCAGCAA 1092
      ::::::::::::::::::::::::::::::::::::::::::::::::::
351 AsnIleArgLysValPheLysAspValArgSerValIleAlaArgTyr 367
      ::::::::::::::::::::::::::::::::::::::::::::::::::
1093 AGCGTCCCGAGCGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1142
      ::::::::::::::::::::::::::::::::::::::::::::::::::
367 rLeuAspGluIleLeuLeuLeu 374
      ::::::::::::::::::::::::::::::::::::::::::::::::::
1143 CTGGAGAGAGATCAACCTGCTG 1164

seq_name: gb_rol:BC005439

seq_documentation_block:
LOCUS BC005439 1974 bp mRNA ROD 03-APR-2001
DEFINITION Mus musculus, guanine nucleotide binding protein, alpha 15, clone
ACCESSION MGC:6007, mRNA, complete cds.
VERSION BC005439
KEYWORDS BC005439.1 GI:13529400
SOURCE MGC.
ORGANISM house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Strausberg, R.
1 (bases 1 to 1974)
Direct Submission
Submitted (27-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: Robert Strausberg, Ph.D.
Tel (301) 496-1350
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC plate: 11 Row: b Column: 17.
Location/Qualifiers
1. 1974
/organism="Mus musculus"
/db_xref="LocusID:14676"
/db_xref="taxon:10090"
/clone="MGC:6007 IMAGE:3592300"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"

```

```

CDS
272..1396
/codon_start=1
/product="guanine nucleotide binding protein, alpha 15"
/protein_id="AAH05439.1"
/db_xref="GI:13529401"
/translation="MARSLTGGCCPMCLTEEEKTAARIDOFINRILLLEOKKORPEELK
LLLLPGESGKSTFIKOMRIHGVGYSEEDRRAPRLLIYONIEVSMOAMIDAMDRLOI
PFSRPSKQHASLAVMTQDPYVSTFEKPYAVAMOXLWRDAGIACVYRRERPHLLDSA
VYVLSHLRISDSYIPTAQVLRSEMPGTGINCYFSVAKTKLRIVDVGCGQSRERK
WHCPENVALIYLASLSEYOCLENDENRNEESLALFSTILELFPWFSTSYILFL
NKTDLIEDKIHSTSLATYFP3FGCPRRDAEAKSFILDMFARVYASCAEPQDQGRKGS
RARRFAHFTCATDTSQVRS3FQVDRSDVSLARYLDEINLL"

BASE COUNT 460 a 380 c 571 g 363 t
ORIGIN

alignment_scores:
Quality: 1705.00 Length: 374
Ratio: 4.658 Gaps: 0
Percent Similarity: 97.861 Percent Identity: 84.759

alignment_block:
US-08-878-801-2 x BC005439
Align seg 1/1 to: BC005439 from: 1 to: 1974

1 MetAlaArgSerLeuThrTyrArgCysCysProTyrCysLeuThrGluAs 17
272 ATGCCCCGGTCCCTGACTTGGGGTGGCTGCTCCATGGTCCCTGACAGAG 321
17 pGluLysAlaAlaAlaArgValAspGlnGluT_eAsnArgIleLeuLeu 34
322 GGAGAGACTCCCGCCAGATCGACCAGAGAGCAACACAGGATTGTTGG 371
34 luGlnLysLysGlnAspArgGlyGluLeuLysIleLeuLeuGlyPro 50
372 AACAGAAAAACAGAGCGCCAGGAAATTCAGAAATCCCTCTGTTGGGGCT 421
51 GlyGluSerGlyLysSerThrPheIleLysGluMetArgIleIleHisG 67
422 GGTGAGAGCGGAGAGTACCTTCATCAAGCAATGCGCATCATTCACGG 471
67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyr 84
472 TGTGGGCTACTCGAGAGGAGGAGCCAGAGCCCTCCGGCTGCTCATCTACC 521
84 lnAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
522 AGAACATCTTCGTCTCCATGCGAGCCCATGATACATGCGATGGACCGCTG 571
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
572 CAGATCCCTTCAGCAGCGCTGACAGCAAGCAACACGCCAGCCTAGTGAT 621
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
622 GACCCAGGAGCCCTATTAAGTGAGCACATTCGAGAAGCCATATCGAGTG 671
134 laMetGlnTyrPheValArgAspAlaGlyIleArgAlaCysTyrGluArg 150
672 CCATGCACTACCTGTCGGGAGCGCGGCATCCGTCATGCTACAGCGA 721
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
722 AGCGTGAAATTCACCTTCGAGCTCCCGGCTGATTAACCTGTCCACACT 771
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
772 GGAGCGCATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 821
184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
822 GCAGTCGATGCCACACAGGAGCATCAATGAGTACTGCTTCTCCGTGAAG 871

```


source		1..1150
		/organism="Xenopus laevis"
		/db_xref="taxon:8355"
		/tissue_type="oocyte"
CDS		39..1118
		/codon_start=1
		/product="guanine nucleotide regulatory protein Gq alpha subunit"
		/protein_id="AA52189.1"
		/db_xref="GI:505690"
		/translation="MTLSIMACCLSEAEKARRINDEIEROLRRDKRDARRELKLL
		LGTSGSGSTFIKMRHIGSGYSDKRGFTKLYQNIPTAMQAMIRAMETLKTIPYK
		YEHNKGHALLVREVDVEKVASFNPIYDAIKYLNWDPGIQECIDRRREYQLSDSKYY
		LNDVRIATQGLYPTQDDYLRVPTTGIIEYFDLSQVIFRMVDVGQRRSRKWIH
		CFENVTSIMFLVALSEYDOVLVESNENMEESKALFRTIITYPWFQSSVILFLNKK
		DILLEKIMYSHLVDYFPEYDGPQDAQAAREFILKMFVDLNPDSDKIITYSHFTCATDT
		ENIRFVAAVKDTILQLNLKEYNLV"
BASE COUNT		351 a 254 c 267 g 278 t
ORIGIN		
alignment_scores:		
Quality:		1057.00 Length: 362
Ratio:		3.571 Gaps: 2
Percent Similarity:		81.768 Percent Identity: 56.354
alignment_block:		
US-08-878-801-2 x XLU10502		..
Align seg 1/1 to: XLU10502 from: 1 to: 1150		
13	CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluIleAs	29
66	TGCCTCAGCGAGGAGGCCAAGGAAGCTCGAAGATTAAACGATGAGATCGA	115
29	nargileLeuLeuGluGlnLysGlnAspArgGlyGluLeuLysLeuL	46
116	GAGGAGCTCCCGCGGGGATAGAGAGAGCCCGCGGGAGCTCAAACTCC	165
46	euleLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet	52
166	TCTGTCTGCGCACAGGAGAAAGTGGAAAAAGCACATTTATCAACAGATG	215
63	ArgIleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr	79
216	AGGATAATCCATGGATCTGGTATTATTCAGATGAAGATAAGAGGGGTTCAC	265
79	gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA	96
266	AAAGCTTGCTATCAAAATATATTCACATGCAATGCAAGCTATGATTCGGG	315
96	laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis	112
316	CCATGGAACCCCTTAAATACCATACCAAGTATGAACACACAAAGGCCAT	365
113	AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGluLy	129
366	GCATGTTAGTTCGAGAGTTGATGTAGAGAGGTGCGCTCTTTTGAGAA	415
129	sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA	146
416	TCCTTATGTAGATGCAATAAAGTATTGTGGAATGACCCCTGGCATACAAG	465
146	laCysTyrGluArgArgArgGluPheHisLeuLeuAspSerAlaValTyr	162
466	AGTGCTACGACGAGCGGAGGAATATCAGCTGCTGATTCCTCAATAATAT	515
163	TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl	179
516	TATCTCAACGACGTGACCGCTATAGCAACTCAAGGATACCTTCCACCCA	565
179	aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC	196
seq_name: gb_ov:XLU10502		
seq_documentation_block:		
LOCUS		XLU10502 1150 bp mRNA VRT 04-NOV-1994
DEFINITION		Xenopus laevis guanine nucleotide regulatory protein Gq alpha subunit mRNA, complete cds.
ACCESSION		U10502
VERSION		U10502.1 GI:505689
KEYWORDS		
SOURCE		African clawed frog.
ORGANISM		Xenopus laevis
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopus laevis
AUTHORS		Shapira, H., Way, J., Lipinsky, D., Oron, Y. and Battey, J.F.
TITLE		Neuroleptin B receptor, expressed in Xenopus laevis oocytes, selectively couples to G alpha q and not G alpha 11 [published erratum appears in FEBS Lett 1994 Aug 1;349(2):318]
JOURNAL		FEBS Lett. 348 (1), 89-92 (1994)
MEDLINE		94298961
REFERENCE		2 (bases 1 to 1150)
AUTHORS		Battey, J.F.
TITLE		Direct Submission
JOURNAL		Submitted (10-JUN-1994) James F. Battey, National Cancer Institute, Laboratory of Biological Chemistry, Bldg. 37, Room 5D-02, Bethesda, MD 20892, USA
FEATURES		Location/Qualifiers

```

566 GCAAGATGCTGAGAGTTCGAGTCCCAACACAGGAGCAATCATTCAGATC 615
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
616 CCTTCGACTTACAAGTGTATCTTCAGAAATGGTGGATGAGGAGACAG 665
213 LysSerGluArgLysThrIleHisCysPheGluAsnValIleAlaLe 229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
666 AGATCTGAGAGAGGAAGATGATACACTGTTTGAACACGTCACTCCAT 715
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluAlaSnA 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
716 AATGTTCCCTGTAGCCCTTAGTGAATATGATCAAGTCTGCTGCAATCTG 765
246 sGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
766 ACAATGAGAACCGAATGGAAGAAAGCAAGCGTTATTTAGGAGCATCATC 815
263 GluLeuProThrPheLysSerThrValIleLeuPheLeuAsnLysThr 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
816 ACATATCCCTGGTGTCAAAATTCCTCCGTCATCCTGTTCTTAACAGAA 865
279 rAspIleLeuGluGlyLysIleProThrSerHisLeuAlaThrTyrPheP 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
866 AGATCTTCTAGAGGAAATCATGCTACTCCGCTCTAGTGTATTTATTC 915
296 rSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
916 CTGATATGATGCGCCCAACCTGATGCAAGCAGCAAGAAATTCATC 965
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValaspGlyProGluG 329
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
966 TGAATATGTTTGTAGATCTC.....ATAATCTACTCGCACTTCACCTGCG 997
329 ySerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
998 CGACAAA.....ATAATCTACTCGCACTTCACCTGCG 1029
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1030 CCACAGATACTGAGAACATTCGTTTCGCTTTCGTCAGTCAAGGACACA 1079
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1080 ATCTACAGCTAAACCTGACGAGTACACCTGGTG 1115

seq_name: gb_ov:XLGTPBIND

seq_documentation_block:
LOCUS      XLGTPBIND      1347 bp      mRNA      VRT      28-APR-1993
DEFINITION Xenopus laevis alpha subunit of Gq Gtp-binding protein mRNA,
            complete cds.
ACCESSION  L05540
VERSION    L05540.1 GI:214187
KEYWORDS   GTP-binding protein.
SOURCE     Xenopus laevis (library: lambda-gtl0 from Doug Melton) ovary cDNA
            to mRNA.
ORGANISM   Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE  Lader, K.J. and Smith, L.D.
            Isolation of the Xenopus cDNA for G alpha q
            Unpublished (1992)
FEATURES   Location/Qualifiers
            1..1347
             /organism="Xenopus laevis"
             /db_xref="taxon:8355"
             /tissue_type="ovary"
             /issue_lib="lambda-gtl0 from Doug Melton"
            1..1080
             /codon_start=1

```

```

/product="GTP-binding protein"
/protein_id="AAA49730.1"
/db_xref="GI:214188"
/translation="MTLESIWACCLSEAEARRINDEIBOLRDKRDARRELKLL
LGTSGSKSTFLIKMRIIHGSCYSDPKRTKLVYQNTFSAMQAMIRAMETLKIPIYK
YEHNGKHALLYREVDEKVA SFENPYDAIKYLMNDPGICEDYDRREYQLSDSKYY
LNDLRIATHGYLPQODLVVRVTTGLIIEFPDQLSVIFRMVDVGQSSRRKWIH
CFENVTSMFLVALSEYQVGVESDNENMEESKALFRTIITYPMFQNSVILFLNKK
DLLEKINYSHLVDYFPYD SPQDQAAREFIILKMFVDLPDSDSKIIYSHTCATDT
ENIRFVAAYKDTILQLNLKSYNLV"
<136..>2156
/bound_moiety="GTP"
1081..>1347
protein_bind      415 a      282 c      306 g      344 t
3'UTR
BASE COUNT
ORIGIN

alignment_scores:
  Quality: 1052.00      Length: 362
  Ratio: 3.566         Gaps: 2
  Percent Similarity: 81.492      Percent Identity: 56.354

alignment_block:
US-08-878-801-2 x XLGTPBIND
Align seg 1/1 to: XLGTPBIND from: 1 to: 1347

13 CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluIleAs 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 TGCCCTCAGCGAGGAGCGCGAGGAACTCGAGCATTTAAGCATGAGATCGA 77
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 rArgIleLeuLeuGluGlnLysLysGlnAspAlaGlyGluLeuLysLeuL 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
78 GAGCGAGCTCCCGCGGATAGAGAGAGCGCCGCCGGGAGCTCAAACTCC 127
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
128 TCCTCTCGGCACAGGAGAAAGTGGAAAAGGCACATTATTCACACAGATG 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 ArgIleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
178 AGGATATCCATGGATCTGGTTATTTCAGATGAAGATAAGAGGGGTTTCA 227
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 AAAGCTTGCTATCAAAATATATTCTCTGCAATGCAAGCTATGATTCGGG 277
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278 CCATGGAACCTTTAAATACCATACAGATGATGAACACACACAGGCCAT 327
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGluLy 129
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 GCACCTGTTAGTTCGAGAAGTTGATGAGAGAAGTCGCGCTCTTTTGAGA 377
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArg 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
378 TCCTTATGATGATCAATAAGATTTTGGAAATGACCCCTGGCATACAG 427
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 laCysTyrGluArgArgGluPheHisLeuLysAspSerAlaValTyr 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
428 AGTGCTACGACGCGGAGGATATCAGCTGTGTGATCCACTAAATAT 477
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 TyrLeuSerHisLeuGluArgIleThrGluGluGluTyrValProThrAl 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 TATCTCAATGACCTGGACCGCTATAGCAACTCATGCTACCTCTCTCCCA 527
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 GCAAGATGCTCAGAGTTCGAGTCCCAACACAGCGGAATCATTCAGTATC 577
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 ySPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
578 CCTTTGACTTACAAACGCGTATCTTCAGAAATGGTGGATGTAGGAGACAG 627
213 LysSerGluArgLysIleHisCysPheGluAsnValIleAlaLe 229
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 CGCTCTGAGAGAGGAAGATGGATACACTGCTTTGAAATGTCACCTCCAT 677
229 uIeTyLeuAlaSerLeuSerGluTyAspGlnCysLeuGluGluAsnA 246
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
678 AATGTTCTCTGTAGCCCTTAGTGAATATGATCAAGTGGTGGATCTG 727
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
728 ACAATGAGAACCGATGGAAGAAAGCAAGCGTTATTAGGAGCATCATC 777
263 GluLeuProTyrPheLysSerThrValIleLeuPheLeuAsnLysTh 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
778 ACATATCCATCGTTTCAAAATCTCAGTCATCCTGTTTCTAAACAAGAA 827
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
828 AGATCTCTAGAGAAATAATCATGTACTGCGCATCTAGTTGATTAATCC 877
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
878 CTGAATATGATGACCCCGAGAGATGCACAGCAGCAGAGAAATTCATC 927
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGlu 329
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
928 TTAAGATGTTCTGATATCTC.....AATCCAGACAG 959
329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
960 CGACAAA.....ATAATCTACTCGCATCTCACCTGTG 991
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
992 CTACAGATACAGAGCAATTCGTTCTGCTGTCAGTCAAGGACACA 1041
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeu 374
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1042 ATCCTACAGCTAAATCTGAAGAGATACAACTGGTG 1077

seq_name: gb_om:DOGGALP

seq_documentation_block: 1077 bp mRNA MAM 02-AUG-1996
LOCUS
DEFINITION Canis familiaris G alpha q mRNA, complete cds.
ACCESSION L76257
VERSION 1
KEYWORDS G alpha q; coupling; thromboxane receptor.
SOURCE Canis familiaris
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 1077)
AUTHORS Dunlop, P.C., Leis, L.A. and Johnson, G.J.
TITLE Canine Thromboxane Receptor-G Protein Coupling
JOURNAL Unpublished (1996)
FEATURES
    source
        1..1077
            /organism="Canis familiaris"
            /db_xref="taxon:9615"
            /cell_type="platelet"
            /dev_stage="adult"
            <1..>1077
                /gene="G alpha q"
            1..1077
                /gene="G alpha q"
            1..>1077
                /gene="G alpha q"
            /codon_start=1
            /protein_id="AAB05548.1"
            /db_xref="GI:1478044"

mRNA
gene
CDS
```

```
/translation="MTLESIMACCLSEAEARRINDEIEROLRRDKRDARREKLLL
LGTGSGSKSTFIKQMRIHSGSYSDERKGTKLQYQIFAMQAMIRAMDTLKIPYK
YEHNKAHAQLYREVVDVEXSAFENPYDAIKLWNPQIOECYDRRREYQLSDSTKY
LNDLORVADPAILYLPQQDYLVRPVPTTGLIEVPTDLOSIFRMVDVGQRRRKH
CFENVTSIMFLVALSEYDOQVLVESNENRMBESKALFTIITYPWFQNSVILFLNKK
DLLEEKIMYSHLVDFEYDPGPDQAQAAREFILKMFVDLPDPSDKIILYSHTCATDT
ENIREVFAAVKDTILQLNLKEYNLV"

BASE COUNT      311 a      259 c      266 g      241 t
ORIGIN

alignment_scores:
    Quality: 1047.00      Length: 362
    Ratio: 3.537      Gaps: 2
    Percent Similarity: 81.768      Percent Identity: 55.801

alignment_block:
US-08-878-801-2 x DOGGALP ..

Align seg 1/1 to: DOGGALP from: 1 to: 1077

13 CysLeuThrGluAspGluLysAlaAlaAArgValAspGlnGluIleAs 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 TGCCTGACGAGAGGCCCAAGAGCCCGCGGATCAACGACGAGATCGA 77
29 nArgIleLeuLeuGluGlnLysGlnAspArgGlyGluLeuLysLeu 46
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
78 GCGGCGAGCTCCGACGAGCAAGCGGAGCGCCCGCGGAGCTCAAGCTGC 127
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
128 TGCCTGCTCGGGACGGGAGAGAGTGGCAAGAGTACGTTTATCAAGCAGATG 177
63 ArgIleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
178 AGAATCATCCATGGTGGCTACTCTGATGAAGACAAAAGGGGCTTCAC 227
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 CAAGCTGGTGTATCAGAACATCTTCAGGCGCATGCGAGGCTATGATCAGAG 277
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278 CCATGGACAGCTCAAGATCCCTACAGTACGAGCAGCAATAAGGCTCAT 327
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGluLys 129
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 GCACAATTAGTTGAGAGAGTCGACGCTGGAGAGGCTGCTGCTTTTGAGAA 377
129 sArgTyrAlaAlaAlaMetGlnTyrLeuTyrArgAspAlaGlyIleArgA 146
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
378 TCCATATGTAGACCGCAATAAGAGTTTATGGAATGATCCTGGAATCCAGS 427
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
428 AATGCTATGACAGACGACGAGAGATATCAGTTGTCGACTCTACCAATAC 477
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 TATCTTAATGACTTGGACCGCTGCTGACCCCGCTTACCTGCTCTAGCA 527
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 ACAAGATGCTTATGAGATTCGAGTCCCGCCACACAGGAGTATCGAATACC 577
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
578 CCTTTGACTTACAAGTGTCTATTTTCAAGATGTCGATGAGGGGCCAG 627
213 LysSerGluArgLysLysTyrPheIleHisCysPheGluAsnValIleAlaLe 229
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 AGGTCAGAGAGAGAAATGGATACACTGCTTTTGAAATATGTCACCTCTAT 677
```


BASE COUNT 309 a 255 c 276 g 240 t
ORIGIN
DLEEKIMVSHLVDPYFPEYDGPQDQAAREFLLKMFVDLNPDSKIIYSHFTCATDT
ENIRFVFAAVKDTILQNLKYNLV"

alignment_scores:
Quality: 1046.00 Length: 362
Ratio: 3.534 Gaps: 2
Percent Similarity: 81.768 Percent Identity: 55.525

alignment_block:
US-08-878-801-2 x AF234260

Align seg 1/1 to: AF234260 from: 1 to: 1080

13 CysLeuThrGluAspGluLysAlaAlaAraValAspGlnGluLeuAs 29
28 TGCCCTGAGCGAGGAGCCCAAGAGAGCCGCGGAGGATCAACGACGATCGA 77
29 nArgileLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeu 46
78 GCGGAGCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 127
46 euLeuLeuGlyProGlyGlySerGlyLysSerThrPheLeuLysGlnMet 62
128 TGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 177
63 ArgileLeuLeuGlyAlaGlyLysSerGluGluGluArgLysGlyPheAr 79
178 AGGATCATCCACGCGGTGGGTGCTCTGATGAGAGAGAGAGGAGGAGGAGGAG 227
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
228 CAACTGGGTATCAGAACATCTTTACAGCCATGCGAGCCATGCTCAGAG 277
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
278 CTATGACACTCTCAAGATCCCATACAAAGTATGAACACAAATAAGGCTCAT 327
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGlyLys 129
328 GCACAAATGGTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 377
129 sArgTyrAlaAlaIaMetGlnTyrLeuTrpArgAspAlaGlyIleArgA 146
378 TCATATATGATGAGCAATTAAGAGGCTTGTGAATGATCTCTGGAATCCAGG 427
146 laCystTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
428 AATGCTACGATAGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 477
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
478 TATCTCAACGACTTGACCGCTGTGGCTGACCTTCTCTATCTGCTACCAATAC 527
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
528 ACAAGATGCTGTAGAGTTCGAGTCCCGCCACAGGAGGAGGAGGAGGAGGAG 577
196 yPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
578 CCTTCGACTTACAGAGTGTATCTTCCAGATGCTGATGATGAGGAGGAGGAG 627
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
628 AGGTCAG 677
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
678 CATGTTCTGCTGAGCGCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 727
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262

246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
247
728 ACATGAGAACCGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 777
263 GluLeuProTyrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
778 ACATACCCCTGCTTCCAGAACTCCCTCGGTTATCTGTCTTTAAACAAGAA 827
279 rAspIleLeuGluGlyLysIleProThrSerHisLeuAlaThrTyrPheP 296
828 AGATCTTCTAGAGAGAGAAATCATCTATCCATCCATCTAGCTGACTACTCC 877
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
878 CAGATATGATGAGCCCGCAGAGAGATGCCAGGAGGAGGAGGAGGAGGAGGAG 927
313 LeuAspMetTyrArgMetTyrThrGlyCysValAspGlyProGluG1 329
928 CTCAAGATGCTCGTGAGCTG.....AATATCTACTCCCACTTCCACGTGCG 991
329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
960 TGACAAA.....ATTATCTACTCCCACTTCCACGTGCG 991
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
992 CCACAGACCGGAGAGATATCCGCTTTGCTTCTGCTGCGCTCAAGGAGAC 1041
363 valLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
1042 ATCTCTCCAGTTGAACCTGAAGGAGTACATCTGCTGTC 1077

seq_name: gb_rol:AF234260

seq_documentation_block: 1080 bp mRNA ROD 26-MAR-2000
LOCUS AF234260
DEFINITION Rattus norvegicus heterotrimeric guanine nucleotide-binding protein
alpha q subunit mRNA, complete cds.

ACCESSION AF234260
VERSION AF234260.1 GI:7329186
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 1080)
AUTHORS Strotmann, R.
TITLE Rat G alpha q subunit
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1080)
AUTHORS Strotmann, R.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2000) Pharmacology, Freie Universitaet Berlin,
Thielallee 69-73, Berlin 14195, Germany

FEATURES
source Location/Qualifiers
1..1080
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="brain"
1..1080
/note="G protein"

CDS
/codon_start=1
/product="heterotrimeric guanine nucleotide-binding
protein alpha q subunit"
/protein_id="AAF5930.1"
/db_xref="GI:7329187"
/translation="MTLESIACCLSEEKARRINDEITERLRDKRDRRELKLL
LGESGKSTFFIKQMRILHSGSVSDKRGFTKLVYONIFTAMQAVRAMDTLKIPYK
YEHKAGQLREVDEKVSAPENPYDAIKSLWNPDPGIECYDRERQVLSDTKYK
LNDLDRVADPSYLPDTPGTGIIETFPDQSVIFRWVDVGORSERRKWIH
CFENVTSIMFLVSEYDQVLVESDNRNREESKALFRLTIITPWFQNSVILFLNKK

```

728 ACAATGAGAACCAAGTGGAGGAGCAAGCACTCTTTAGAACCATATC 777
263 GluLeuProThrPheLysThrSerValIleLeuPheLeuAsnLysTh 279
778 ACATATCCCTGGTTCAGAACTCCTCTGTATTCTCTTAAACAAGAA 827
279 rAspIleLeuGluGlyIleProThrSerHisLeuAlaThrTyrPhe 296
828 ACATCTCTAGAGGAGAAATATGATTCCACCTAGTCGACTACTTCC 877
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
878 CAGATATATGAGCCAGAGAGATGCCAGGACGACGAGAAATTCATC 927
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGlu 329
928 CTGAAGATGTCGTGGACCTG.....AACCCGACAG 959
329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCys 346
960 TGACAAA.....ATCATCTACTCGCACTTCAGGTGTG 991
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgaspSer 362
992 CCACAGACAGAGAACATCCGCTCGTGTTCCTGCTGCTCAAGGACACC 1041
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
1042 ATCTGTCAGCTGAACCTGAGGAGTACATCTGGTC.1077

```

seq_name: gb_om:BOVGL2A

```

seq_documentation_block: 1391 bp mRNA MAM 07-FEB-1999
LOCUS BOVGL2A
DEFINITION Bovine mRNA for GTP-binding protein alpha-subunit (GL2A), complete
ACCESSION D90336
VERSION 1
KEYWORDS G protein; GL2-alpha; GTP-binding protein; GTPase.
SOURCE Bovine liver, cDNA to mRNA, clone pGL7.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1391)
Nakamura,F., Ogata,K., Shiozaki,K., Kameyama,K., Ohara,K., Haga,T.
and Nakada,T.
Identification of two novel GTP-binding protein alpha-subunits that
lack apparent ADP-ribosylation sites for pertussis toxin
J. Biol. Chem. 266 (19), 12676-12681 (1991)
91286303
These data kindly submitted in computer readable form by: Fumio
Nakamura
Department of Biochemistry
Institute of Brain Research
Faculty of Medicine, University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113
Japan
phone: 03-3812-2111 x3561
Fax: 03-3814-8154.
Location/Qualifiers
1..1391
/organism="Bos taurus"
/db_xref="taxon:9913"
251..1312
/feature="GTP-binding protein, alpha-subunit"
/codon_start=1
/protein_id="BAA14350.1"
/db_xref="GI:217568"
/translation="MACCLSDVYKSRINAEIKLRDRKDRARRELKLLLLTGTS
GKSTFIKQMRHIIHGAGYSEEDKRGFTKLIVQNIPTAQAMIRAMETIKILYEQNKA

```

FEATURES

source

CDS

```

NALLREVDVVEKVTTEHRYVSAIKTLWNPQIOBCYDRREYQLSDSAKYYLTVDVR
IATSGYLPQDDVLVRVPVTGLIEYFDLENIIFRMVDVGGQSRERRWJHCFENVT
SIMELVASEYDQVLVESDNIINRMEEKALKERTIVTFWPNSSVILFNKRDLEDK
ILHSHLVDFPFGGRRDPAKAREFILKMFVDLNPDSKLIYSHTCATDTENIRFV
FAAVKDTILOLNLAKEYNLV" 212 t
BASE COUNT 282 a 426 c 471 g
ORIGIN

```

alignment_scores:

Quality: 1046.00 Length: 362
Ratio: 3.570 Gaps: 2
Percent Similarity: 80.939 Percent Identity: 56.077

alignment_block:

US-08-878-801-2 x BOVGL2A ..

Align seg 1/1 to: BOVGL2A from: 1 to: 1391

```

13 CysLeuThrGluAspGluLysAlaAlaArgValAspGlnGluIleAs 29
|||||..... ||| ..... |||||..... |||||.....
260 TGCCTGAGCGATGAGGTGAAGGAGTCCAAGCGGATCAACGCCGAGATCGA 309
29 nArgIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeuL 46
||||| ..... ||| ..... ||| ..... |||||.....
310 GAAGCAGCTGCGCGGCGGACAAAGCGCGCGCGGAGCTCAAGCTGC 359
46 euleLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 52
|||||..... ||| ..... ||| ..... |||||.....
360 TCGTCTCGCGCGCGGAGAGCGGAGAGCAAGGAGCAAGTTCATCAAGCAGATG 409
63 ArgIleIleHisGlyValArgLysSerGluGluGluArgLysGlyPheAr 79
|||||..... ||| ..... ||| ..... |||||.....
410 CGCATCATCCAGCGCGGCGGTACTCAGAGAGGACAAAGCGGCGGTTCAC 459
79 gProLeuValThrGlnAsnIlePheValSerMe:ArgAlaMetIleGluA 96
|||||..... ||| ..... ||| ..... |||||.....
460 CAAACTGCTGTACCAAGACATCTTACCCTCCATCAGGCCATGATCCGCG 509
96 laMetGluArgLeuGlnIleProPheSerArgP:ogLysSerLysHis 112
|||||..... ||| ..... ||| ..... |||||.....
510 CCATGGAACCTGAAGATCTCTCAAGTACGAGCAGCAAGCAAGGCCAAC 559
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129
|||||..... ||| ..... ||| ..... |||||.....
560 GCATCTGTATCCGCGGCGGTGGATGTGGAAGAGTGACGACCTTCGAGCA 609
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArg 146
|||||..... ||| ..... ||| ..... |||||.....
610 CCGTACGTGAGCGCCATCAAGACCTGTGGAAAGAGCCCGCGCATCCAGG 659
146 laCystTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
|||||..... ||| ..... ||| ..... |||||.....
660 AGTGCTATGACCGCGCGGAGTACCAGCTCTCGACTCCGCCAAGTAC 709
163 TyrLeuSerHisLeuGluArgIleThrGluGluLysTyrValProThrAl 179
|||||..... ||| ..... ||| ..... |||||.....
710 TACCTGACGCGCGTGGACCGCATTCACCTCAGCTACCTGCCACCCA 759
179 agLAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
|||||..... ||| ..... ||| ..... |||||.....
760 GCAGGAGCTGCTCGCGGTGCGCGTCCACCGCGGCGGCGGCGGCGGCG 809
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
|||||..... ||| ..... ||| ..... |||||.....
810 CCTTCACCTGGAGAACATCATCTCAGGATGGTGTGTGGGGGGCGAG 859
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
|||||..... ||| ..... ||| ..... |||||.....
860 AGGTCCGCGGAGGAGTGGATTCACTGCTTTGAGAACGTCAGCTCCAT 909
229 ulletYrLeuAlaSerLeuSerGluTyrAspGlnLysLeuGluGluAsnA 246
|||||..... ||| ..... ||| ..... |||||.....

```

LNLD R VAD P A Y L P T Q Q D V L R V R P T T G I L E Y P F D L Q S V I I
C F E N V T S I M F L V A L S E Y D Q V L V E S D N E N R M E E S K A L F R T I I

LND

246 sngInGluAsnArgMeLLysGluSerLeuhtatdca.hscf

```

948 ACAATGAGAACCGAATGAGGAAAGCAAGGCTCTCTTTAGAACAAATTATC 997
263 GluLeuProTrrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
998 ACATACCCCTGGTTCCAGAACTCCTCGTTATTCTGTCTTAAACAAGAA 1047
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
1048 AGATCTTCTAGAGGAGAAATCATGTATTCCCATCTAGTCGACTACTTCC 1097
296 rOSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
1098 CAGATATGATGGACCCAGAGAGATGCCGAGCGAGCGCGAGAAATTCATT 1147
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
1148 CTGAAGATGTCTGTGGACCTG.....AACCAGACAG 1179
329 ySerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
1180 TGACAAAATTAC.....TACTCCCACTTCAGGTGCG 1211
346 lathrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
1212 CCACACACACCGAGAAATCCGCTTTGCTTTGCTGCCGCTCAAGGACACC 1261
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
1262 ATCTCTCAGTTGAACCTGAAGAGGTACAAATCTGGTC 1297
```


OM of: US-08-878-801-2 to: N_Geneseq_0601.* out_format : pfs
Date: Sep 6, 2001 1:04 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/6004808/runat_06092001_110135_4516/app_query.fasta_1.437
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=p2n.rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -XGAPEXT=0.500
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=6004808 -CGWL_1_293 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT_THREADS=1

Search information block:
Query: US-08-878-801-2
Query length: 374
Database: N_Geneseq_0601.*
Database sequences: 730101
Database length: 313950809
Search time (sec): 77.720000

score_list:	Strd Orig	zScore	EScore	Len	Documentation
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAA48751		1360.00	3777.99	1.8e-202	2060
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV03464		1347.00	3759.41	2.0e-201	1125
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAV03463		1711.00	3300.56	7.1e-176	1125
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV03465		1701.00	3281.12	8.7e-175	1125
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31056		1037.00	1990.60	6.6e-103	1077
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31058		1037.00	1976.54	4.0e-102	3831
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31060		1037.00	1975.56	4.5e-102	4185
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31065		1037.00	1975.44	4.6e-102	4231
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31062		1037.00	1975.37	4.7e-102	4257
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31064		1035.50	1973.23	6.1e-102	3969
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31063		1033.50	1969.51	9.9e-102	3909
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31093		1033.00	1982.79	1.8e-102	1080
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ19461		1033.00	1982.79	1.8e-102	1080
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ37399		1013.00	1940.24	4.2e-100	1503
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV54416		746.50	1420.70	3.7e-71	1703
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ49048		746.50	1420.70	3.7e-71	1703
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31790		745.00	1421.89	3.1e-71	1402
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ10854		745.50	1418.13	5.1e-71	1802
/SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ51098		742.50	1412.93	9.9e-71	1703
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ09737		740.00	1408.32	1.8e-70	1664
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ15502		738.50	1406.24	2.3e-70	1543
/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ69004		725.50	1376.67	1.0e-68	2274
/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT:AAZ41294		719.00	1364.22	5.1e-68	2236
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ57661		707.00	1348.29	3.9e-67	1146
/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ69002		692.50	1314.44	3.0e-65	1910
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ53006		668.00	1266.37	1.4e-62	1987
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ49451		651.00	1236.31	6.8e-61	1516
/SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT:AAZ14527		643.00	1220.46	5.2e-60	1558
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ41972		638.00	1210.50	1.9e-59	1591
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAZ62461		601.50	1132.85	4.0e-55	2310
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAZ62460		601.50	1132.63	4.1e-55	2967
/SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT:AAZ58272		598.00	1128.97	6.5e-55	2235
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ92615		526.50	994.66	2.0e-47	1461
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ36330		520.50	981.56	1.1e-46	1662
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ47419		520.50	981.56	1.1e-46	1662
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ13598		435.00	825.36	5.3e-38	672
/SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT:AAZ66090		424.00	809.72	3.9e-37	400
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ69735		380.00	723.09	2.6e-32	441
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ76649		360.00	675.43	1.2e-29	974
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ16974		354.00	665.95	4.0e-29	799

seq_name: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAA48751
seq_documentation_block:
ID AAA48751 standard; cDNA; 2060 BP.
XX
AC AAA48751;
XX
DT 08-SEP-2000 (first entry)
XX
DE Human G-alpha-16 nucleotide sequence.
XX
KW Human; G-alpha-16; G protein; cytostatic; hyperproliferative disorder;
KW cancer; inflammation; infection; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 220..1344
FT /*tag= a
FT /product= "G-alpha-16"
XX
PN WO200032817-A1.
XX
PD 08-JUN-2000.
XX
PF 25-AUG-1999; 99WO-US19613.
XX
PR 03-DEC-1998; 98US-0205143.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Cowser LM;
XX
DR WPI; 2000-412354/35.
XX
P-PSDB; AAY99841.
XX
PT A new antisense compound for inhibiting the expression of human
PT G-alpha-16 and treating, preventing or delaying infections,
PT inflammation or hyperproliferative disorders such as cancer -
XX
PS Example 13; Page 82-84; 100pp; English.
XX
CC The present sequence encodes the human G-protein G-alpha-16.
CC G-alpha-16 interacts differentially with several receptor types
CC including members of the opioid and chemokine receptor families.
CC Antisense oligonucleotides that target the present sequence can be
CC used to modulate the expression of G-alpha-16. They may be used to
CC inhibit the expression of G-alpha-16 in human cells and tissues and thus
CC to treat diseases associated with G-alpha-16, such as hyperproliferative
CC disorders, especially cancer. Infections, inflammation or tumour
CC formation can be prevented or delayed. The compounds can be
XX used in research and diagnostics in sandwich and other assays.
SQ Sequence 2060 BP; 415 A; 631 C; 640 G; 374 T; 0 other;

alignment_scores:
Quality: 1960.00 Length: 374
Ratio: 5.241 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-878-801-2 x AAA48751 ..
Align seg 1/1 to: AAA48751 from: 1 to: 2060
1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17

1120 CCTAAGCAGGATGCTGAGGCACCAAGAGGTT⁺ATCTGGACATGTACAC 1169

317 rArgMetTyrThrGLyCysValAspGLyProGluGLySerLysLysGlyA 334
|||||
1170 GAGGATGTACACCGGGTCTGTGAGCGCCCGAGGGCAGCAAGAGGGCG 1219

334 laArgSerArgLeuPheSerHisTyrThr⁺ysAlaThrAspThrGln 350
|||||
1220 CACGATCCGACGCGCTTTTCAGCCACTACACA⁺GTGCACAGACACACAG 1269

351 AsnIleArgLysValPheLysAspValArgAspSerValIleuAlaArgTy 367
|||||
1270 AACATCCGCAAGTCTTCAAGGACGTCGGGA⁺TCGGTGCTGCCCGCTA 1319

367 rLeuAspGluIleAsnLeuLeu 374
|||||
1320 CTGGACGAGATCAACCTGCTG 1341

```
seq_documentation_block:
```

AAV03464;

DT 22-JUN-1998 (first entry)

Promiscuous G-alpha 16 protein; mouse; G-protein coupled receptor;
KW
ligand; agonist; antagonist; signal transduction; ss
KW

XX
PN
W09748820-A1

24-DEC-1997.

PF 20-JUN-1997;

PR 21-JUN-1996;

PA (AURO-) AURORA
XXXX
XX

DR P-PSDB; AAW424

PT Stable cells

XX
FI and agonists of

XX
XX

16 protein (see

normally coupled

protein. A mi-

may also compr

G-alpha protein

of signal transduction.

receptor of the
permits charac

Sequence 1125

```

alignment_scores:
  Quality: 1947.00      Length: 374
  Ratio: 5.206          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 99.198

alignment_block:
  US-08-878-801-2 x AAV03464 ..
  Align seg 1/1 to: AAV03464 from: 1 to: 1125

1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
|||||
1 ATGGCCCGCTGCTGACCTGGCGCTGCTGCCCTGGTGGCTGACGAGGA 50
|||||
17 pGluLysAlaAlaArqValAspGlnGluIleAsnArqIleLeuLeuG 34
|||||
51 TGAGAAAGCCGCCCGGGGTGGACCGAGGATCAACAGATCTCTTGG 100
|||||
34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
|||||
101 ASCAGAAGAAGCAGACCGCGGGAGCTGAAGTGTGCTTTTGGGCCCA 150
|||||
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
|||||
151 GGCAGAGCGGGAAGACACCTTCATCAAGACAGATCGGATCATCCACGG 200
|||||
67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
201 CCGCGGCTACTGGAGAGAGCGCAAGGCCCTTCGGGCCCTGGTCTACC 250
|||||
84 lnaSnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
251 AGAACATCTTCGTGTCATCGCGGCCATGATCGAGGCCATGGAGCGGTG 300
|||||
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
|||||
301 CAGATTCCATTACAGACGCCGAGAGCAGCAGCACCACCGCTAGCCTGGT 350
|||||
117 tSerGlnAspProTyrLysValThrThrPheGluLysArgTyrAlaAla 134
|||||
351 GAGCCAGGACCCCTATAAAGTGACCCAGCTTCGAGAAGCGTACGCTGGG 400
|||||
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
401 CCATGAGTGGCTGTGGAGGATGCGCGCATCCGGGGCTGCTATGAGCGT 450
|||||
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 CGCGGGGAATTCACCTGCTCGATTACAGCCGTGTACTACCTGTCCCACCT 500
|||||
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnaspValLeuA 184
|||||
501 GGAGCGCATCACCGAGGAGGCTACGTCGCCACACAGCTCAGACAGTGCCT 550
|||||
184 rgSerArgMetProThrThrClyIleasnGluTyrCysPheSerValGln 200
|||||
551 GCAGCCGCATGCCACCACCTGGCATCAACGAGTACTGCTTCTCCGTCGAG 600
|||||
201 LysThrAsnLeuArgIleValAspValGlyGlyGlnLysSerGluArgly 217
|||||
601 AAAACCAACCTCCGGATCGTGGACGTCGCGGGCCAGAGTCAAGCGGTAA 650
|||||
217 sLysTrpIleHisCysPheGluasnValIleAlaLeuIleTyrLeuAla 234
|||||
651 GAAATGGATCCATTGTTTCGAGAACGTGATCGCCCTCATCTACCTGGCCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnAsnGlnGluasnArg 250
|||||
701 CACTGAGTGAATACGACCGAGTCCCTGGAGGAGAACACACGAGGAAACCG 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267

```


296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3629 CAGAATATGATGGACCCAGAGAGATGCCAGCGACCGGAGAAATTCATT 3678
 313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG1 329
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3679 CTGAAGATGTTGGTGAGACTG.....AACCCAGACAG 3710
 329 ySerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3711 TGACAAA.....ATTATCTACTCCACCTACGCTGG 3742
 346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3743 CCACAGACACCGAGAAATATCGCTTTGCTTGTGCTGCTCAAGGACACC 3792
 363 ValLeuAlaArgTyrLeuAspGluIleLeu 373
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3793 ATCTCCAGTTGAACCTGAAGGACTGCGGTCTG 3825

seq_name: /SIDSI/gcdata/geneseq/geneseq/NA1999.DAT:AA231060

seq_documentation_block:

ID AA231060 standard; cDNA; 4185 BP.

AC AA231060;

DT 07-JAN-2000 (first entry)

XX pmGluR2/Car*Galphaq15 fusion construct cDNA sequence.

XX G-protein fusion receptor; Car; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder; ss.

XX Homo sapiens.

XX WO9951641-A1.

XX 14-OCT-1999.

XX 02-APR-1999; 99WO-US07333.

XX 03-APR-1998; 98US-0080671.

XX (NPSF-) NPS PHARM INC.

XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simlin RT;

XX WPI; 1999-610995/52.

XX P-P5DB; AAY49129.

PT New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease -

PS Example 1; Fig 11; 255pp; English.

XX The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a Car
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (Gq) linked to ICD or the
 CC linker. (I), and recombinant chimeric receptors (CR) without the Gp
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or

CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used:
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of Car and mGluR
 CC domains allows presentation of GABABR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.

XX Sequence 4185 BP; 891 A; 1253 C; 1150 G; 891 T; 0 other;

alignment_scores:

Quality: 1037.00 Length: 361

Ratio: 3.515 Gaps: 2

Percent Similarity: 81.717 Percent Identity: 55.402

alignment_block:

US-08-878-801-2 x AA231060 ..

Align seg 1/1 to: AA231060 from: 1 to: 4185

13 CysLeuThrGluAspGluLysAlaAlaAArg/valAspGlnGluIleAs 29
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3133 TGCCTGAGCGAGGAGGCCAAGCGCGGATCAACGACGAGATCGA 3182
 29 nArgIleLeuLeuGluLysLysGlnAspArg/GlyGluLeuLysLeu 46
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3183 CGCGCAGCTCGCGAGGACACAGCGGCGCGCGGCGGAGCTCAAGCTGC 3232
 46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3233 TGCCTGCGGACGAGGAGAGTGGCAGAGTATGTTTATCAACGACGATG 3282
 63 ArgIleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3283 AGAATCATCTCATGGTCAAGTACTCTGATGAAATATAAGGGGCTTCAC 3332
 79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3333 CAAGCTGGTGTATCAGAACATCTTTCAGCGGCAT(CAGGCCATGATCAG 3382
 96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3383 CCATGGACACACTCAAGATCCATACCAAGTATG/GCACAAATAGGCTCAT 3432
 113 AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGluLy 129
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3433 GCACAAATAGTTCGAGAGTGTATGCGAGAGCTGTCTCTTTTGAGAA 3482
 129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArg 146
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3483 TCCATATGTAGATGCAATAAGAGATTATGGAATGATCCTCGAATCCAGG 3532
 146 laCystTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3533 AATGCTATGATACGACGAGAAATATCAATATCTGACTCTACCAATAC 3582
 163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3583 TATCTTTAATGACTTGGACCGCGTAGCTGACCCCTGCTACCTGCTACGCA 3632
 179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3633 ACAAGATGTGCTTAGAGTTCGAGTCCGCCACACACACATCATCGAATACC 3682
 196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGluIn 212
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 3683 CCTTTGACTTACAAAGTGTCTATTTTTCAGAAATGCTGATGAGGGGCCAA 3732

WPI: 1999-6100995/52.
P-PSDB; AAY49134.

New G-protein fusion receptors and chimeras containing domains from different receptors, used to screen for modulators, potentially useful e.g. for treating or preventing stroke or Alzheimer's disease -

Example 1; Fig 11; 255pp; English.

The invention relates to G-protein fusion receptors (I) comprising:

(1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and intracellular (ICD) domains, each chosen independently from a CaR (calcium receptor), GluR (metabotropic glutamate receptor) and GABAAR (gamma-aminobutyric acid receptor); (2) an optional linker attached to the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the linker. (I), and recombinant chimeric receptors (CR) without the GP component, are used to assess function of the various domains and to identify compounds (e.g. allosteric modulators or antagonists) that act on these domains. The modulators are potentially useful for treating or preventing diseases associated with the receptors, e.g. stroke, head or spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia, Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive disorders and depression. Nucleic acid (II) that encodes (I) is used: (1) for recombinant production of corresponding proteins; and (2) to produce cells used in screening for modulators. Use of CaR and mGluR domains allows presentation of GABAAR domains, to a binding agent, in a form more like the natural domain structure compared with use of incomplete receptors, lacking one or more domains. By shuffling different domains, agents can be identified that affect particular domains of a receptor.

Sequence 4231 RP: 894 A: 1275 C: 1161 G: 901 T: 0 other;

seq_documentation_block:	
IID	AAZ31065 standard; cDNA; 4231 BP.
XX	
XX	AAZ31065;
XX	
XX	07-JAN-2000 (first entry)
XX	
XX	pmGluR2/Car*galphai15+3ala linker fusion construct cDNA sequence.
XX	
XX	G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
XX	metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
XX	gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
XX	spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
XX	Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
XX	cognitive disorder; ss.
XX	
XX	Homo sapiens.
XX	
XX	WO9951641-Al.
XX	
XX	14-OCT-1999.
XX	
XX	02-APR-1999; 99WO-US07333.
XX	
XX	03-APR-1998; 98US-0080671.
XX	
XX	(NPSP-) NPS PHARM INC.
XX	
XX	Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
XX	Simin RT;
XX	

```

Align seg 1/1 to: AA231065 from: 1 to: 4231

13 CysLeuThrGluAspGluLysAlaAlaArgValAspGlnGluLeAs 29
|||||:||||: ||| |||:||||:|||||:CG
3142 TGCCTGCGAGGAGGCCAAGGAGCCGGCGGATCAACGACGAGCAGA 3191

29 nArgIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeuL 46
||||| ||| :||||:||||: ||| |||:||||:|||||:
3192 GCGCAGCTCCGACGGGACACAGCGGAGCCCGCGGGAGCTCAAGCTGC 3241

46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
||||| ||| :||||:||||: ||| |||:||||:|||||:
3242 TGCTGCTCGGCACAGGACGAGAGTGGCAAGAGTACGTTTATCAAGCAGATG 3291

63 ArgIleIleHisGlyAlaGlyTyrSerGluGluGluAtrGlyLysPheAr 79
||||| ||| :||||:||||: ||| |||:||||:|||||:
3292 AGAATCATCCATGGCTGAGGATCTCTGATGAAGATAAAGGGCTTCAC 3341

79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
||||| ||| :||||:||||: ||| |||:||||:|||||:
3342 CAAGCTGGTGTATCAGACATCTTTCAGGCCCATGAGGCCATGATCAGAG 3391

96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
||||| ||| :||||:||||: ||| |||:||||:|||||:
3392 CCATGGACACACTCAAGATCCCATACAGTATGACGACACAATAGGCTCAT 3441

113 AlaSerLeuValMetSerClnAspProTyrLysValThrThrPheGluLy 129
||||| ||| :||||:||||: ||| |||:||||:|||||:
3442 GCACAAATTTAGTTCGAGAGAGTTCATGGGAGAGGCTGCTGCTTTTGAGAA 3491

129 sArgTyrAlaAlaAlaMetGlnTriLeuTriPArgAspAlaGlyIleArg 146

```

```

3492 10CATATGATGATCAATAAAGAGTTTATGGAATGATCTCTGAATCCAG 3541
146 10CysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
3542 10ATGCTATGATAGACGACGAATATCAATATCTGACTCTACCAATAC 3591
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
3592 10ATCTTAATGACTTGGACCGGTAGCTACCCCTGCTACCTGCTACGCA 3641
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
3642 ACAAGATGCTTGAAGTTCAGTCCGCCACACAGGATCATCGAATACC 3691
196 yspSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
3692 CCTTTGACTTCAAGTGTCTTTCAGAAATGTCGATGAGGGGCCAA 3741
213 LysSerGluArgLysTyrTrpIleHisCysPheGluAsnValIleAlaLe 229
3742 AGGTCAGACAGACAAATGATGATACCTGCTTTGAAATGTCACCTCTAT 3791
229 uileTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
3792 CATGTTCTAGTACGCTTGTAGTATGATGATGATGATGATGATGATGAT 3841
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
3842 ACAATGAGAACCGAATGAGAAAGAGGCTCTCTTTAGAACCAATATC 3891
263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
3892 ACATACCCCTGTTTCCAGAACTCTCGGTATTCTGTTCTTTAAACAAGAA 3941
279 rAspIleLeuGluLysLysThrProThrSerHisLeuAlaThrTyrPheP 296
3942 AGATCTCTAGAGAGAGAAATCATGTTATCCATCTAGTCTGACTTCTC 3991
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
3992 CAGAATATGATGAGCCACAGAGATGCCAGGAGCCGCGAGATTCATT 4041
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
4042 CTGAAGATGTTCTGGACCTG.....ATTATCTACTCCACTTCACGTGCG 4073
329 ySerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
4074 TGACAA.....ATTATCTACTCCACTTCACGTGCG 4105
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
4106 CCACAGACCGAGAAATATCGCTTTGCTCTGCTCCGCTCAAGGACACC 4155
363 valLeuAlaArgTyrLeuAspGluIleAsnLeu 373
4156 ATCTCTCAGTTGAACCTGAAGGACTCGGCTCTG 4188

```

seq_name: /SIDS1/gcgdata/genesecq/genesecq/NA1999.DAT:AAZ31062

seq_documentation_block:
ID_AA231062 standard; cDNA; 4257 BP.

XX AA231062;

XX 07-JAN-2000 (first entry)

DE mGluR8/Car*Galphai5 fusion construct cDNA sequence.

XX G-protein fusion receptor; Car; calcium receptor; GluR; head injury;
KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;

```

KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;  
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;  
XX cognitive disorder; ss.  
OS Homo sapiens.  
XX WO9951641-A1.  
XX 14-OCT-1999.  
XX 02-APR-1999; 99WO-US07333.  
XX 03-APR-1998; 98US-0080671.  
XX (NPSP-) NPS PHARM INC.  
XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
PI Simin RT;  
XX WPI; 1999-610995/52.  
XX P-PSDB; AAY49131.  
XX New G-protein fusion receptors and chimeras containing domains from  
XX different receptors, used to screen for modulators, potentially useful  
XX e.g. for treating or preventing stroke or Alzheimer's disease -  
XX Example 1; Fig 11; 255pp; English.  
XX The invention relates to G-protein fusion receptors (I) comprising:  
XX (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and  
XX intracellular (ICD) domains, each chosen independently from a Car  
XX (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR  
XX (gamma-aminobutyric acid receptor); (2) an optional linker attached to  
XX the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the  
XX linker. (I), and recombinant chimeric receptors (CR), without the GP  
XX component, are used to assess function of the various domains and to  
XX identify compounds (e.g. allosteric modulators or antagonists) that act  
XX on these domains. The modulators are potentially useful for treating or  
XX preventing diseases associated with the receptors, e.g. stroke, head or  
XX spinal cord injury, epilepsy, ischaemia, hypoglycaemia, cognitive  
XX disorders and depression. Nucleic acid (II) that encodes (I) is used;  
XX (1) for recombinant production of corresponding proteins; and (2) to  
XX produce cells used in screening for modulators. Use of Car and mGluR  
XX domains allows presentation of GABABR domains, to a binding agent, in a  
XX form more like the natural domain structure compared with use of  
XX incomplete receptors, lacking one or more domains. By shuffling different  
XX domains, agents can be identified that affect particular domains of a  
XX receptor.  
XX Sequence 4257 BP; 1152 A; 1044 C; 1046 G; 1015 T; 0 other;

```

alignment_scores:

Quality: 1037.00 Length: 361
Ratio: 3.515 Gaps: 2
Percent Similarity: 81.717 Percent Identity: 55.402

alignment_block:

US-08-878-801-2 x AA231062

Align seg 1/1 to: AA231062 from: 1 to: 4257

```

13 CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluIleAs 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3205 TGCTGAGCGAGGAGGCCAAGGAGCGCGCGGATCAACGACGATCGA 3254
29 nArgIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLeuLeu 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3255 GCGGACGCTCCGCGAGGACGAGCGGAGCGCGCGCGCGCGCGCGCG 3304
46 eutLeuLeuGlyProGlyGluSerGlyLysSerTrpPheIleLysGlnMet 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Thu Sep 6 15:16:52 2001

```
3305 TGCTGCTGGGACAGGAGAGAGTGGCAAGAGTAGCTTTATCAAGCAGATG 3354
63 ArgilelleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
|||||
3355 AGAATCATCCATGGTCCAGATATCTCTGATGAGATAAAGGGGCTTCAC 3404
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
|||||
3405 CAAGCTGGTGTATCAGAACATCTTCCAGGCCATGCAGGCCATGATCAG 3454
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
|||||
3455 CCATGGACACACTCAAGATCCATACAGATATGAGCACATTAAGGCTCAT 3504
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129
|||||
3505 GCACANTTAGTTCGAGAGAGTGTGAGGAGAGGTGTCTGCTTTTGAGAA 3554
129 sArgTyrAlaAlaLaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
|||||
3555 TCCATATGATGATCAATAAAGAGTGTGGAATGATCCTGGAATCCAGG 3604
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
|||||
3605 AATGCTATGATAGACGACGAGATATCAATATCTGACTCTACCAAAATAC 3654
163 TyrLeuSerHisGluArgIleThrGluGluGlyTyrValProThrAl 179
|||||
3655 TATCTTAATGACTTGGACCGCTAGTACCTGCTGCTACCTGCTTACGCA 3704
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
|||||
3705 ACAAGATGTCTAGATGTTGAGTCCACACAGGATCATCGAATACC 3754
196 yspPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
|||||
3755 CCTTGACTTACAAAGTGTCTTTTTCAGAAATGCTGATAGGGGCCAA 3804
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
|||||
3805 AGGTGACGAGAGAGAAATGGATACACTGCTTTGAAATGTCACTCTAT 3854
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
|||||
3855 CATGTTCTAGTACGCTAGTGAATATGATCAAGTCTCTGCTGGAGTCA 3904
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
|||||
3905 ACAATGAGAACCAATGGAGGAAGCAAGGCTCTCTTTAGAACAAATATC 3954
263 GluLeuProThrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
|||||
3955 ACATACCCCTGGTTCACAAACTCCTCGTTATTTCTGTTCTTAACAAGAA 4004
279 rAspIleLeuGluGlyLysIleProThrSerHisLeuAlaThrTyrPheP 296
|||||
4005 AGATCTTCTAGAGAGAAATCATGTTATCCCATCTAGTCCGACTACTTCC 4054
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaLysArgPheIle 312
|||||
4055 CAGAATATGATGGACCCAGCAGAGATCCCGAGGCGAGGAGGAAATCAT 4104
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluGl 329
|||||
4105 CTGAAGATGTTCTGGGACCTG.....AATCTACTCCCACTTACGTCG 4136
329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
|||||
4137 TGACAAA.....ATTATCTACTCCCACTTACGTCG 4168
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
|||||
4169 CCACAGACACCGAGAAATATCGCTTTGCTTGTGCGGTCAGGACACC 4218
```

```
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
|||||
4219 ATCTCTCCAGCTGAACCTGAGACCTGCGGCTCG 4251
```

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31064

seq_documentation_block:

ID AAZ31064 standard; cDNA: 3969 BP.

XX AAZ31064;

XX 07-JAN-2000 (first entry)

XX GABA-BR1a*Gqo5 fusion construct cDNA sequence.

XX G-protein fusion receptor; CaR: calcium receptor; GluR: head injury;
KW metabotropic glutamate receptor; GABABR: chimeric receptor; stroke;
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW cognitive disorder; ss.

XX Homo sapiens.

XX WO9951641-A1.

XX 14-OCT-1999.

XX 02-APR-1999; 99WO-US07333.

XX 03-APR-1998; 98US-0080671.

XX (NPSP-) NPS PHARM INC.

XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;

PI Shimin RT;

XX WPI; 1999-610995/52.

XX P-PSDB; AAY49133.

New G-protein fusion receptors and chimeras containing domains from
different receptors, used to screen for modulators, potentially useful
e.g. for treating or preventing stroke or Alzheimer's disease -

XX Disclosure; Fig 13; 255pp; English.

XX The invention relates to G-protein fusion receptors (I) comprising:
CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC intracellular (ICD) domains, each chosen independently from a CaR
CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC linker. (I), and recombinant chimeric receptors (CR) without the GP
CC component, are used to assess function of the various domains and to
CC identify compounds (e.g. allosteric modulators or antagonists) that act
CC on these domains. The modulators are potentially useful for treating or
CC preventing diseases associated with the receptors, e.g. stroke, head or
CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC disorders and depression. Nucleic acid (II) that encodes (I) is used:
CC (1) for recombinant production of corresponding proteins; and (2) to
CC produce cells used in screening for modulators. Use of CaR and mGluR
CC domains allows presentation of GABABR domains, to a binding agent, in a
CC form more like the natural domain structure compared with use of
CC incomplete receptors, lacking one or more domains. By shuffling different
CC domains, agents can be identified that affect particular domains of a
CC receptor.

XX Sequence 3969 BP; 945 A; 1079 C; 1065 G; 880 T; 0 other;

alignment_scores:

Quality: 1035.50 Length: 361
 Ratio: 3.510 Gaps: 3
 Percent Similarity: 81.717 Percent Identity: 55.402
 alignment_block:
 US-08-878-801-2 x AA231064 ..
 Align seg 1/1 to: AA231064 from: 1 to: 3969
 13 CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluLeuAs 29
 2920 TGCTGTAGCGAGGAGGCAAGGAGGAGGCGCGGATCAACGACGAGATCGA 2969
 29 nArgIleLeuLeuGluGlnLysGlnAspArgGlyGluLeuLysLeuL 46
 2970 GCGGCGAGCTCCGCGAGGAGCAAGCGGAGCGCGCGGAGCTCAAGCTGC 3019
 46 euLeuLeuGlyProGlyGlySerGlyLysSerThrPheIleLysGlnMet 62
 3020 TGTGTCTGGGACGAGAGAGTGGCAAGAGTACGTTTATCAAGCAGATG 3069
 63 ArgIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
 3070 AGAATCATCTGGTGGTACGATCTCTGATGAAGATAAAGGGCTTCAC 3119
 79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
 3120 CAAGCTGGTGTATCAGAACATCTTCAGGCCATGAGCGCATGATCAGAG 3169
 96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
 3170 CCATGGACACACTCAAGATCCCTCAAGTATGAGCACAATAAGGCTCAT 3219
 113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129
 3220 GCACATAGTTCGAGAGTGTGATGGAGAGGTGTCTGCTTTTGAGAA 3269
 129 sArgTyrAlaAlaLaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
 3270 TCCATGTAGATGATGATAAAGAGTGTATGAAGTATGATCTGGAATCCAGG 3319
 146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
 3320 AATGCTATGATGAGCAGGAGCAATATCAATATCTGACTTACCAAAATAC 3369
 163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
 3370 TATCTTATGACTTGGACCGGTAGCTGACCTGCCTACCTACGCA 3419
 179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
 3420 ACAAGATGTGCTTAGAGTTCGAGTCCCGCCACACAGGATCATCGAATAC 3469
 196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
 3470 CCTTTGACTTACAAAGTGTATTTTCAGATGTCGATGAGGGGCCAA 3519
 213 LysSerGluArgLysLysTyrPheHisCysPheGluAsnValIleAlaLe 229
 3520 AGGTACAGAGAAGAAATGATACACTGCTTGAATAATGTCACCTCTAT 3569
 229 uileTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluCysAsnA 246
 3570 CAGTTTCTAGTAGCGCTTAAAGTAAATATGATCAAGTCTCTCGTAGAGTAC 3619
 246 snGlnGluAsnArgMetLysGlnLeuAlaLeuPheGlyThrIleLeu 262
 3620 ACAATGAGACCAATGAGGAAGCAAGCTCTCTTTAGAACAAATATC 3669
 263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
 3670 ACATACCCCTGGTTCAGGAACCTCCTCGGTATCTCTTCTTAAACAAGA 3719

279 rAspIleLeuGluGlyLysIleProThrSerHisLeuAlaThrTyrPheP 296
 3720 AGATCTTCTAGAGGAGAAATCATGTATTCCTAGTACTGACTACTTCC 3769
 296 roSerPheGlnGlyProLysGlnAspAlaGluAlaLysArgPheIle 312
 3770 CAGNATATGATGGACCCAGAGAGATGCCAGCAGCCCGAGAAATTCATT 3819
 313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
 3820 CTGAAGATGTTCTGGACCTG.....AACCCAGACAG 3851
 329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
 3852 TGCAAA.....ATTATCTCTCCACTTCACGTGC 3883
 346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
 3884 CCACAGACACCGAGAAATATCCGCTTGTCTGTCGCGTCAAGGACACC 3933
 363 ValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
 3934 ATCTCT.....CAGTTGAACCTG 3951
 seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31063
 seq_documentation_block:
 ID AAZ31063 standard; cDNA; 3909 BP.
 XX AAZ31063;
 XX 07-JAN-2000 (first entry)
 XX GABA-BR2*Gqo5 fusion construct cDNA sequence.
 XX G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABAAR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder; ss.
 OS Homo sapiens.
 XX WO951641-A1.
 XX 14-OCT-1999.
 XX 02-APR-1999; 99MO-US07333.
 XX 03-APR-1998; 98US-0080671.
 XX (NPSF-) NPS PHARM INC.
 XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 XX WPI; 1999-610995/52.
 DR P-PSDB; AAY49132.
 XX New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease
 XX Disclosure; Fig 13; 255pp; English.
 XX The invention relates to G-protein fusion receptors (1) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a CaR
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABAAR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
 CC linker. (1), and recombinant chimeric receptors (CR) without the GP

100

100

XX	Homo sapiens.
XX	Location/Qualifiers
XX	Key 1..1080
FT	CDS /tag= a
FT	/product= "G-alpha-11"
XX	US951455-A.
XX	PI Cowsert LM;
XX	PD 14-SEP-1999.
XX	P-PSDB; AAY29789.
XX	PP 04-DEC-1998; 98US-0205922.
XX	PR 04-DEC-1998; 98US-0205922.
XX	PA (ISIS-) ISIS PHARM INC.
XX	PI WPI; 1999-539140/45.
XX	DR P-PSDB; AAY29789.
XX	DT Inhibitory antisenase compounds useful for the treatment of diseases
XX	PT associated with G-alpha-11
XX	PS Example 13; Column 45-48; 38pp; English.
XX	CC The present invention describes inhibitory antisenase compounds of 8-30
XX	CC nucleotides, targeted to a nucleic acid molecule encoding human
XX	CC G-alpha-11. The present sequence encodes human G-alpha-11. AAZ19468 to
XX	CC AAZ19547 represent human G-alpha-11 phosphorothioate antisense
XX	CC oligonucleotides given in the present invention. The oligonucleotides
XX	CC may be useful for the treatment of diseases associated with G-alpha-11.
XX	CC Sequence 1080 BP; 245 A; 334 C; 314 G; 186 T; 0 other;
XX	SQ
alignment_scores:	
Quality: 1033.00 Length: 362	
Ratio: 3.550 Gaps: 2	
Percent Similarity: 80.387 Percent Identity: 55.525	
alignment_block:	
US-08-878-801-2 x AAZ19461 ..	
Align seg 1/1 to: AAZ19461 from: 1 to: 1080	
13	CysLeuThrGluAspGlnLysAlaAlaArgValAspGlnGluIleAs 29
28	TGCCTGGCGATGAGTCAAGGAGTCACAAAGCGGATCAACGCCGAGATCGA 77
29	nArgIleLeuGluGlnLysLysGlnAspArgGlyLeuLysLeuL 46
78	GAAAGAGCTGCGCGGGAACAAGCGGACCCCGCGGAGCTCAAGCTGC 127
46	eutLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
128	TGCTGTCTGCGCAGCGGAGAGCGGAAAGAGCACGTTTCATCAAGCATG 177
63	ArgIleIleHisGlyNalaglyTyrrSerGluGluArgLysGlyPheAr 79
178	CGCATCATCCAGCGCGCGCTACTCGGAGGAGGACAGCGCGGCTTAC 227
79	gProLeuValTyrrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
228	CAGCTCGTCTACCAGAACATCTTACGCCCATGAGCCCATGATCCGGG 277
96	laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
278	CCATGGAGAGCTCAAGATCTCTACAGTAGTACGACGAGAACAGGCCAAT 327
113	AlaSerLeuValMetSerGlnAspProTyrrLysValThrThrPheGluLy 129

Thu Sep 6 15:16:52 2001

us-08-878-801-2.p2n.rng

Page 18

372 nLeu 373
1169 GCTC 1172


```
|||||
351 GAGCCAGACCCCTATAAAGTGACCGCTTTGAGAACGCTACCTCGG 400
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
401 CCATGCAGTGGCTGTGGAGGATGCCGGCATCCGGGCTGCTATGACGCT 450
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 CGCGGGGAATTCCACCTGCTCGATTACGCGTGTACTACCTGTCCACCT 500
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
501 GGAGCGCATCACCGAGGAGGCTACGTCCTCCACAGCTCAGACGCTGCTC 550
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
551 GCAGCGCATGCCACCATCGGCATCAACAGGACTGCTTCTCGGTGCAG 600
201 LysThrAsnLeuArgIleValAspValGlyClnLysSerGluArgLy 217
|||||
601 AAAACCACTGGGATCGTGGAGCTCGGGGCCAGAGTCAGAGCGTAA 650
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
|||||
651 GAAATGATCCATGTTTCGAGAACGTGATCGCCCTCATCTACCTGGCCT 700
234 erLeuSerGluTyrAspGlnCysLeuGluAsnGlnAsnArg 250
|||||
701 CACTGAGTAAACGACCGAGTGGCTGGAGGAGAACACAGGAGAACCGC 750
251 MeTysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProtrpPh 267
|||||
751 ATGAAAGAGAGCTCCGATGTTTGGAGCTATCTGGAACCTACCTGGT 800
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGlu 284
|||||
801 CAAAGACATCCGCTCATCTCTTCTCAACAAACCGACATCTCGAGG 850
284 lLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
851 AGAAATCCCACTCCACCTGGCTGACCTATTTCCACAGTTCACAGGC 900
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
901 CCTAAGCAGGATGCTGAGGAGCAGCAAGAGTTTCATCTGGACATGTAC 950
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
|||||
951 GAGGATGTACACCGGCTGCTGGACGCCCGGAGGAGCAGCAAGAGGCG 1000
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1001 CAGATCCCGACGCTTTTCAGCCACTACATGTGCCACAGACACAG 1050
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
|||||
1051 ACATCCGAAGGCTTCAAGGAGCTGGGGAGCTCGGTGCTCGCCGCTA 1100
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1101 CTTGGACGAGATCAACTGCTG 1122
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-09-205-143-1

```
seq_documentation_block:
; Sequence 1, Application US/09205143
; Patent No. 6107091
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF G-ALPHA-16 EXPRESSION
; FILE REFERENCE: RTS-0032
; CURRENT APPLICATION NUMBER: US/09/205,143
```

; CURRENT FILING DATE: 1998-12-03

; NUMBER OF SEQ ID NOS: 87

; SEQ ID NO 1

; LENGTH: 2060

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (220)..(1344)

; US-09-205-143-1

alignment_scores:

Quality: 1960.00

Ratio: 5.241

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-08-878-801-2 x US-09-205-143-1

Align seg 1/1 to: US-09-205-143-1 from: 1 to: 2060

```
1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
|||||
220 ATGCCCGCTCGCTGACCTGGCTGCTGCCCTGGTGCCTGACGAGGA 269
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeu 34
|||||
270 TGAGAGGGCGCGCGCGGTGGACAGGAGATCAACAGGATCCTCTTGG 319
34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
|||||
320 AGCAGAAGAAGACGAGCGCGGGAGCTGAAGCTGCTGCTTTGGGCCCA 369
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
|||||
370 GCGCAGAGCGGGAAGACACCTTCATCAAGCAGATGCGGATCATCCACG 419
67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
420 CGCCGCTACTCTCGAGGAGGAGGCAAGGCTTCGGGCCCTGGTCTACC 469
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
470 AGAACATCTTCGTGTCATCGCGGCCATGATGAGGCCATGGAGCGCTG 519
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
|||||
520 CAGATTCATTCACAGCGCCGAGAGCAAGCACACGCTAGCCTGGTCTAT 569
117 tSerGlnAspProTyrLysValThrThrPheGlyLysArgTyrAlaAla 134
|||||
570 GAGCCAGGACCCCTATAAGTGACACGTTTGA3AAGCGCTACGCTGCGG 619
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
620 CCATGAGTGGCTGTGGAGGATGCCGGCATCCGGCCTGCTATGAGCGT 669
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
670 CGCGGGGAATTCCACCTGCTCGATTACGCGTGTACTACCTGTCCACCT 719
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
720 GGAGCGCATCACCGAGGAGGCTACGTCCTCCACAGCTCAGACGCTGCTC 769
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
770 GCAGCGCATGCCACCATCGGCATCAACAGGACTGCTTCTCGGTGCAG 819
201 LysThrAsnLeuArgIleValAspValGlyGlyLysLysSerGluArgLy 217
|||||
820 ANAACCACTCGGATCGTGGAGCTCGGGGCGAGAAAGTCAGAGCGCTAA 869
```

```
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
|||||
870 GAAATGGATCCATTTGTTGGAGACGTATGCCCTCATCTACCTGGCT 919
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
920 CACTGAGTGAATACGACCAAGTCCCTGGAGGAGAACACAGGAGAACGC 969
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
|||||
970 ATCAAGGAGAGAGCTCGCATTTGTTGGGACTATCTGGAACTACCTGGTT 1019
|||||
267 elySerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
1020 CAAAAGCACATCGTCATCTCTTCTCAACAAACCGACATCTGGAGG 1069
|||||
284 luLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
1070 AGAAATCCCCACCTCCACCTGGCTACCTATTTCCTCCAGTTTCCAGGCG 1119
|||||
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
1120 CTTAAGCAGAGTCTGAGGAGCCAGAGAGTTTCATCTCGACATGTACAC 1169
|||||
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysGlyA 334
|||||
1170 GAGATGTACACCGGTGCTGGAGCGCCCGGAGGCGCAGCAAGAGGGCG 1219
|||||
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1220 CACGATCCGACGCCCTTTTCAGCCACTACACATGTGCCACAGACACAG 1269
|||||
351 AsnIleArgLysValPheLysAspValArgAspSerValIleAlaArgTy 367
|||||
1270 AACATCCGAAGGTCTTCAAGGACGTGCGGGACTCGGTGCTGCCCGCTA 1319
|||||
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1320 CTTGGAGAGATCAACCTGCTG 1341
```

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-878-801-3

```
seq_documentation_block:
; Sequence 3, Application US/08878801
; Patent No. 6004808
; GENERAL INFORMATION:
; APPLICANT: Negulescu, Paul
; APPLICANT: Offermanns, Stefan
; APPLICANT: Simon, Melvin
; APPLICANT: Zuker, Charles
; TITLE OF INVENTION: PROMISCUOUS G-PROTEINS COMPOSITIONS AND THEIR USE
; FILE REFERENCE: 08366/002001
; CURRENT APPLICATION NUMBER: US/08/878,801
; CURRENT FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: US 60/020,234
; EARLIER FILING DATE: 1996-06-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1122)
US-08-878-801-3
```

alignment_scores:
Quality: 1705.00 Length: 374
Ratio: 4.658 Gaps: 0
Percent Similarity: 97.861 Percent Identity: 84.759

```
alignment_block:
US-08-878-801-2 x US-08-878-801-3
Align seg 1/1 to: US-08-878-801-3 from: 1 to: 1125

1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
|||||
1 ATGCCCGGCTCCCTGACTTGGGCTGCTCCCTGGTGCCTGACAGAGGA 50
|||||
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeu 34
|||||
51 GGAGAAAGACTGCCCCCAAGATCGACAGGAGATCAACAGGATTTGTGG 100
|||||
34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
|||||
101 AACAGAAAAAACAAGAGCGCGAGGAATTGAAACTCTCTGCTGGGCGCT 150
|||||
51 GlyGluSerGlyLysSerThrPheIleIysGlnMetArgIleIleHisG 67
|||||
151 GGTGAGAGCGGGAAGAGTACGTTTCATCAAGCAGATGCGCATTCACGG 200
|||||
67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyr 84
|||||
201 TGTGGGTACTCGGAGGAGGACCGCAGAGCCTTCGGGCTGCTCATCTAC 250
|||||
84 lnAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
251 AGAATCTTCTGCTCCATGTCAGGCCATGATAGATGCGATGGACCGGCTG 300
|||||
101 GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValMet 117
|||||
301 CAGATCCCTTTCAGCAGGCCCTGACAGCAAGCAGCGCCAGCCTAGTAGT 350
|||||
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
|||||
351 GACCCAGGAGCCCTATATAAGTGAGCACCATTTCAGAGGCCATATGCA 400
|||||
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
401 CCATGCGACTACTGTGGGGAGCGGCGCATCCGTCGTCATGTCAGCGGA 450
|||||
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 AGGCGTGAATTCACCTTCTGGACTCCGCGGTGTTATTCCTGTGCACCT 500
|||||
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeu 184
|||||
501 GGAGCGCATATCAGAGGACAGCTACATCCCACCTGCGCAAGCGTGTGC 550
|||||
184 rSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
551 GCAGTCCGATGCCACCAACAGGCATCAATGAGTACTGCTCTCTCCGTGA 600
|||||
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgGly 217
|||||
601 AAACACAAACTCGCATCTGGATGTTGGTGGCCAGAGGTGACAGCGGTAG 650
|||||
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
|||||
651 GAAATGGATTCATCTGTTTCGAGAACGTGATTCCTCATCTACCTGGCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
701 CCCTGAGCGAGTATGACCAAGTGCCTAGAGGAGAGAGATCAGGAGAACGC 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
|||||
751 ATGGAGGAGAGTCTCGCTCTGTTTCAGCAGATCTCTAGAGCTGCCCTGGT 800
|||||
267 elySerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CAAGAGCACCTCGGTCATCTCTTCTCAACAGAGCGGACATCTCTGGAAG 850
|||||
```

```

284 luLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
851 ATAAGATTACACCTCCACCTCCACCTGCCACATATCTCCACAGCTCCAGGGA 900
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
901 CCCCGGCGAGACGAGAGCGCGCAAGAGCTTCATCTTGGACATGTATGC 950
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
951 GCGCGTGTACGCGAGCTGCGCAGAGCCCGAGCGGTGGCAGGAAGGCT 1000
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
1001 CCCGCGCGCGCGCTTCTTCCGACACTTCACCTGTGCCAGGACACGCAA 1050
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
1051 AGCGTCGCGAGCGTGTCAAGAGCTGCGGAGCTCGGTGCTGGCCCGGTA 1100
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
367 rLeuAspGluIleAsnLeuLeu 374
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
1101 CTTGACGAGATCAACCTGTGT 1122
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||

```

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-205-922-1

```

seq_documentation_block:
; Sequence 1, Application US/09205922
; Patent No. 5951455
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF G-APLHA-11 EXPRESSION
; FILE REFERENCE: RTS-0030
; CURRENT APPLICATION NUMBER: US/09/205,922
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 1
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1080)
US-09-205-922-1

```

alignment_scores:
 Quality: 1033.00 Length: 362
 Ratio: 3.950 Gaps: 2
 Percent Similarity: 80.387 Percent Identity: 55.525

alignment_block:

```

US-08-878-801-2 x US-09-205-922-1 ..
Align seg 1/1 to: US-09-205-922-1 from: 1 to: 1080

13 CysLeuThrGluAspGluLysAlaAlaArgValAspGlnGluIleAs 29
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
28 TGCTGAGCGATGAGGTGAAGGAGTCCAAAGCGATCAACGCCGAGATCGA 77
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
29 nArgIleLeuGluGlnLysLysGlnAspArgGlyGluLeuLeuL 46
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
78 GAAGCAGCTGCGCGCGGACAAAGCGAGCGCCCGCGGAGCTCAAGCTGC 127
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
128 TGCTGCTCGCGCGGCGGAGAGCGGAGAGACGTTTCATCAAGACGATG 177
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
63 ArgIleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
178 CGCATCATCAGCGCGCGGCTACTCTCGGAGGAGGACAAAGCGCGGCTTAC 227
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||

```

```

79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
228 CAAGCTCGTCTACCAAGAACATCTTCAACCGCATGCAGGCCCATGATCCGGG 277
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
96 laMetGluArgLeuGlnIleProPheSerArgIletoGluSerLysHisHis 112
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
278 CCATGGAGACGCTCAAGATCTCTACAAGTACACGACACAACAGCCCAAT 327
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGluL 129
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
328 GCGTCTCTGATCCGGAGGTGAGCTGGAGAACGTGACCACTCTCGAGCA 377
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpAlaGaspAlaGlyIleArgA 146
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
378 TCAGTACGTCAAGTCCATCAAGACCTCTGTGGGTGAGACCCGGGCATCCAGG 427
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
428 AATGCTACGACCGCGAGCGCGAGTACCAAGCTCTCCGACTCTGCCAAGTAC 477
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
163 TyrLeuSerHisLeuGluArgIleThrGluGlyTyrValProThrAl 179
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
478 TACCTGACCGACGTGACCGCATCGCCACCTTGGCTACCTGCCACCCA 527
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
179 aglnAspValLeuArgSerArgMetProThrTrpGlyIleAsnGluTyrC 196
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
528 GCAGGAGCTGCTGCGGGTCCCGCTGCCACCAACCGGCATCATCGATACC 577
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
578 CTTTCGACCTGGAGAACATCATCTTCCGGATGGTGGTGGGGGCGAG 627
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
213 LysSerGluArgLysLysTyrIleHisCysPheGluAsnValIleAlaLe 229
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
628 CGGTGCGGCGGAGGAGTGGATCCACTGCTTTCAGAACGTGACATCCAT 677
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluAsnA 246
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
678 CATGTTCTCTGTCGCTCAGCAATACGACCAAGTCTGTGGAGTGG 727
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
728 ACAACGAGAACCGGATGGAGGAGAGCAAGCCCTGTTCGGGACCATCATC 777
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
778 ACCTACCCCTGGTTCAGAACTCCTCGTCATCTCTTCTCAACAAGAA 827
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
828 GGACCTGCTGGAGGACAAAGATCTCTACTCGCACTGCTGGACTACTTCC 877
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaLysArgPheIle 312
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
878 CCGAGTTTCGATGTTCCCGACGCGGACGCCAGGCGCGGGAGTTCATC 927
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
928 CCGAAGATGTTCTGGACCTG.....AACCCCGACAG 959
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
329 ysSerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
960 CGAAG.....ATCATCTA-TCACACTTCACGTGTG 991
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
992 CCACCGACACGAGGAGAACATCCGCTTCGTGTTCGCGCGCGGAGTTCAC 1041
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
1042 ATCTCTGAGCTGAACCTGAAGGAGTACAACTCTGTCTC 1077
    ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||

```

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-07-868-353A-11


```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-407-804-20
seq_documentation_block:
; Sequence 20, Application US/08407804
; Patent No. 5817759
; GENERAL INFORMATION:
; APPLICANT: Margolskee, Robert F.
; TITLE OF INVENTION: Gustducin Materials and Methods
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,804
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/045,801
; FILING DATE:
; APPLICATION NUMBER: US 07/868/353
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5817759and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 114...1175
; US-08-407-804-20

alignment_scores:
Quality: 746.50 Length: 368
Ratio: 2.893 Gaps: 4
Percent Similarity: 70.109 Percent Identity: 42.391

alignment_block:
US-08-878-801-2 x US-08-407-804-20 ..
Align seg 1/1 to: US-08-407-804-20 from: 1 to: 1703
14 LeuThrGluAspGluLysAlaAlaAArgValAspGlnGluLeuAsnAr 30
126 ATTAGTTCAGAGCAGGAGGAGTCCAGCCAAAGGTCCTCAAGAACTGGAGAA 175
30 gileLeuLeuGluGlnLysGlnAspArgGlyLeuLeuLeuLeuL 47
176 GAAGCTTCAGAGATGCTGACACAGATGCAAGAACTGGAAGTTGCTGC 225
47 euLeuGlyProGlyGluSerGlyLysThrPheLeuLysGlnMetArg 63
226 TATTAGGACAGGGAATCAGGAAAGTACTATTTGTTAAACAAATGAAG 275

```

```

64 IleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheArgPr 80
276 ATCATCCCAAGAANTGGTTACGTAACAAGATGTCATGGAGTTTAAAGC 325
80 oLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluAla 97
326 AGTGGTTTACAGTAACACAGTTCAGTCCATCCCGGCCATTGTGAAGCCA 375
97 etGluArgLeuGlnIleProPheSerArgProIleSerLysHisHisAla 113
376 TGACTACACTAGGATGATTATGTCATCCGGAAGTAGAGAGGACCAA 425
114 SerLeuValMetSerGlnAspProTyrLysValThrThrPheGlu.... 128
426 CAACTGCTTCTCTCCATG.....GCAAACACACTAGAAAGATGG 463
129 .....LysArgTyrAlaAlaAlaMetG..nTrrPLeuTrrArgAspA 142
464 TGACATGACGCCCTCAGTTGGCTGAATAATAAACGTCCTGTGGGGCGATC 513
142 laGlyIleArgAlaCysTyrGluArgArgIlePheHisLeuLeuAsp 158
514 CAGGAATTCAGCCTGCTTCGAAGGGCATCTAAATACCAGCTCAATGAC 563
159 SerAlaValTyrTyrLeuSerHisLeuGluArgIleThrGluGluGly 175
564 TCTGCAGCTTACTACCTTATGACTTAGATAGACTCACAGCCCTGGGTA 613
175 rValProThrAlaGlnAspValLeuArgSerArgMetProThrThrGlyI 192
614 TGTGCCAAATGAACAAGACGTTTACATTCCTCCGGTGAACACCTGGTA 663
192 leAsnGluTyrCysPheSerValGlnLysThrAsnLeuArgIleValAsp 208
664 TCATTGAACTCAATTCCTCTTAAAGACTTGAACTTCAGAAATGTTGAT 713
209 ValGlyGlnLysSerGluArgLysLysTrpIleHisCysPheGluAs 225
714 GTAGTGGCCAGAGATCAGAAAGAAAGAAATGATCCACTGCTTTGAAGG 763
225 nValIleAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysL 242
764 AGTGACCTGCATTATATTTTGTGTCAGCCCTAAATGCTCCATCATGAT 813
242 euGluGluAsnAsnGlnGluAsnArgMetLysIleSerLeuAlaLeuPhe 258
814 TTGTAGAAGATGAAGAGGTGAACAGAAATGCATAAAGTCTTCACTCTTC 863
259 GlyThrIleLeuGluLeuProTrrPheLysSerThrSerValIleLeuPh 275
864 AACAGCATCTGTAATACAGATATTTTGCACACACCTCCATTGTTCTGTT 913
275 eLeuAsnLysThrAspIleLeuGluLysIleProThrSerHisLeuA 292
914 TCTTAACAAGAAAGATCTCTTCCAGAGAGAAAGTACCAAGGTGCACCTCA 963
292 laThrTyrPheProSerPheGlnGlyProLysIleAspAlaGluAlaAla 308
964 GCATCTGTTTCCCAAGATACACTGGACCAATAATTCGAAGATGCAGGG 1013
309 .....LysArgPheIleLeuAspMetTyrThrArgMetTyrThrGl 322
1014 AACTACATCAAGAACCAGTCTCTAGACCTGAATCTTAAATA..... 1052
322 yCysValAspGlyProGluGlySerLysGlyAlaArgSerArgArgL 339
1053 .....AAGAAGATAAGGAAA 1068
339 euPheSerHisTyrThrCysAlaThrAspThrIleAsnIleArgLysVal 355
1069 TCTATTCTCACATGCACCTGCCTACTACACACAATAAACGTCATAATCTGT 1118

```



```
alignment_scores:
  Quality: 740.00
  Ratio: 2.879
  Percent Similarity: 70.219
  Length: 366
  Gaps: 4
  Percent Identity: 42.896

alignment_block:
```



```
271 TGGAGCGCTAAAGATTGACTTTGGGAAGCTCCAGGCGAGATGATGCC 320
111 HishisAlaSerLeuValMetSerGlnAspProTyrLysValThrPh 127
: : : : : : : : : : : : : : : : : : : : : : : : : :
321 CGCAATTATTGTTTGTAGCTGGCAGTGCTGAAGAAGAGTCATGATCC 370
: : : : : : : : : : : : : : : : : : : : : : : : : :
127 eGluLysArgTyrAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyI 144
: : : : : : : : : : : : : : : : : : : : : : : : : :
371 AGAA.....CTAGCAGGAGTGAATAAACGGTTATGGCGAGATGGTGGG 414
: : : : : : : : : : : : : : : : : : : : : : : : : :
144 leargAlaCysTyrGluArgArgGluPheHisLeuLeuAspSerAla 160
: : : : : : : : : : : : : : : : : : : : : : : : : :
415 TACAAGCTTCTTCAGCAGATCCAGGAATATCAGCTCAATGATTCGTG 464
: : : : : : : : : : : : : : : : : : : : : : : : : :
161 ValTyrLysSerHisLeuGluArgIleThrGluGluGlyTyrValPr 177
: : : : : : : : : : : : : : : : : : : : : : : : : :
465 TCATATTATCTAATGATCTGGATAGATATCCAGTCTCACTACATCC 514
: : : : : : : : : : : : : : : : : : : : : : : : : :
177 oThrAlaGlnAspValLeuArgSerArgMetProThrThrGlyIleAsn 194
: : : : : : : : : : : : : : : : : : : : : : : : : :
515 AACTCAGCAAGATGTTCTCGGACGAGAGTGAAGACCACAGGCATTGTAG 564
: : : : : : : : : : : : : : : : : : : : : : : : : :
194 lutyCysPheSerValGlnLysThrAsnLeuArgIleValAspValGly 210
: : : : : : : : : : : : : : : : : : : : : : : : : :
565 AAACATATTTCACCTTCACAGACTATATCTCAAGATGTTGTATGATAGT 614
: : : : : : : : : : : : : : : : : : : : : : : : : :
211 GlyGlnLysSerGluArgLysLysTyrIleHisCysPheGluAsnValI 227
: : : : : : : : : : : : : : : : : : : : : : : : : :
615 GGCCAAAGATCAGAACCAAAAAGTGGATTCACGCTTTTGAGGGAGTGAC 664
: : : : : : : : : : : : : : : : : : : : : : : : : :
227 eAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGlu 244
: : : : : : : : : : : : : : : : : : : : : : : : : :
665 AGCAATTATCTTCGTGTGGCCCTCAGTATTATGACCTTGTCTGTGCTG 714
: : : : : : : : : : : : : : : : : : : : : : : : : :
244 luasnAsnGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThr 260
: : : : : : : : : : : : : : : : : : : : : : : : : :
715 AGGACGAGGAGATGACGAATGATGAAAGCATGAAACTGTTTGCACAGC 764
: : : : : : : : : : : : : : : : : : : : : : : : : :
261 IleLeuGluLeuProTyrPheLysSerThrSerValIleLeuPheLeuAs 277
: : : : : : : : : : : : : : : : : : : : : : : : : :
765 ATTGTGTAATAACAATGGTTTACAGAAGCTCAATCATCTCTCCCTTAA 814
: : : : : : : : : : : : : : : : : : : : : : : : : :
277 nLysThrAspIleLeuGluGluLysIleProThrSerHisLeuAlaThr 294
: : : : : : : : : : : : : : : : : : : : : : : : : :
815 CAAGAAAGACCTTTTTCAGGAAAAAATAAGAGGAGTCCGTTAACTATCT 864
: : : : : : : : : : : : : : : : : : : : : : : : : :
294 yrPheProSerPheGlnGlyProLysGlnAspAlaGluAlaLysArg 310
: : : : : : : : : : : : : : : : : : : : : : : : : :
865 GTTATCCAGAATACACAGTTCCTCAATATGAAAGAGGAGCT..... 908
: : : : : : : : : : : : : : : : : : : : : : : : : :
311 PheIleLeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyPr 327
: : : : : : : : : : : : : : : : : : : : : : : : : :
909 .....GCCTATATTCAATGCCAGTTT..... 929
: : : : : : : : : : : : : : : : : : : : : : : : : :
327 oGluGlySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyr 344
: : : : : : : : : : : : : : : : : : : : : : : : : :
930 .GAAGATCTGAACAGAGAAAGATACCAAGGAGATCTATCTACATCA 978
: : : : : : : : : : : : : : : : : : : : : : : : : :
344 hrCysAlaThrAspThrGlnAsnIleArgLysValPheLysAspValArg 360
: : : : : : : : : : : : : : : : : : : : : : : : : :
979 CCTGTGCACAGACACGCAAGAAATGTGCAGTTTGTGTTGATGCTGTACA 1028
: : : : : : : : : : : : : : : : : : : : : : : : : :
361 AspSerValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
: : : : : : : : : : : : : : : : : : : : : : : : : :
1029 GATGTCATCATTAACAACACTTAAAGGAATGTGGACTT 1067
: : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:us-08-371-930-28
seq_documentation_block:
; Sequence 28, Application US/08371930
; Patent No. 5578451
; GENERAL INFORMATION:
; APPLICANT: Nishimoto, Ikuo
```

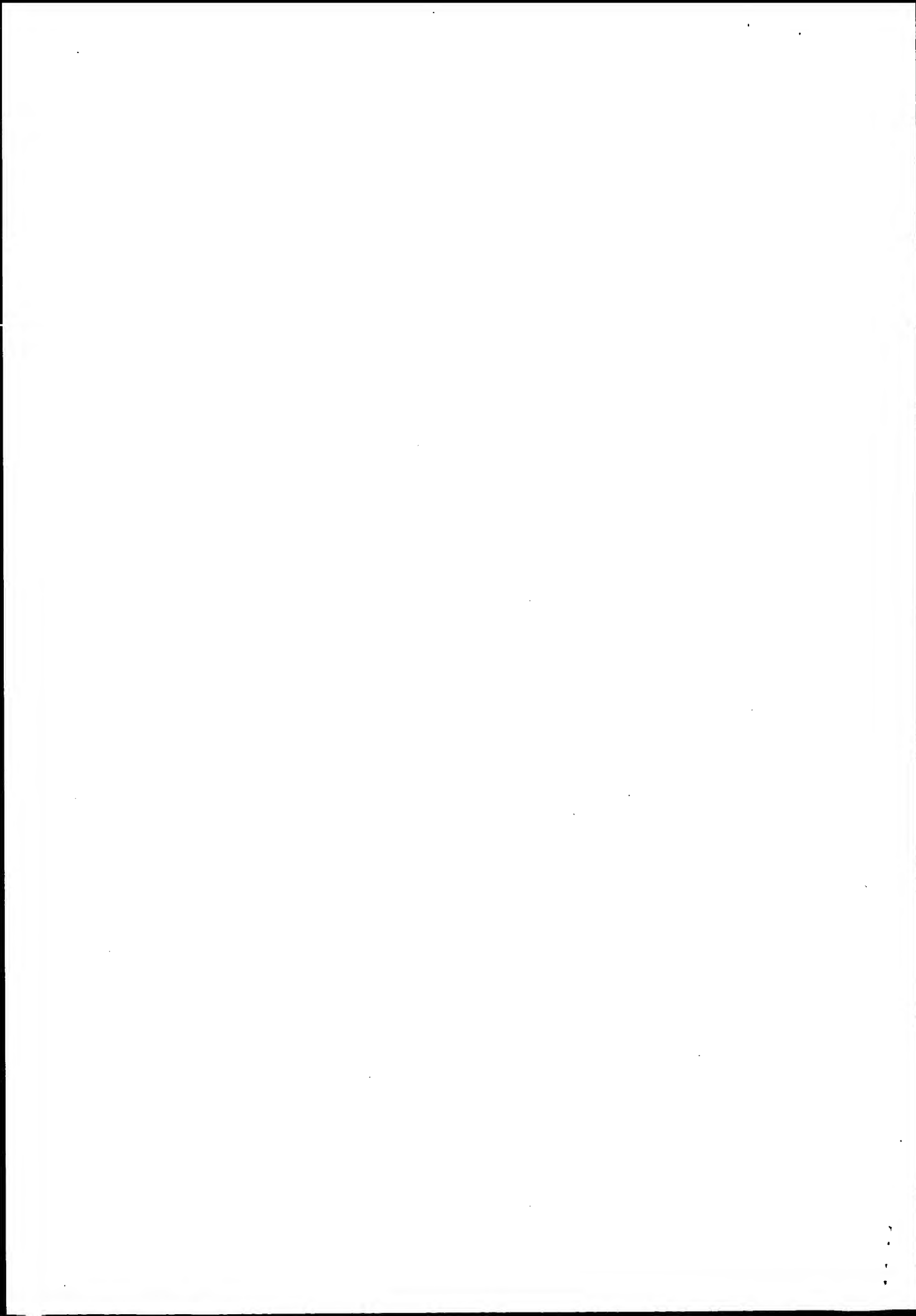
```

; TITLE OF INVENTION: ALZHEIMER'S DISEASE THERAPEUTICS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/371,930
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/019,208
; FILING DATE: February 18, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/154001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2274
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-371-930-28

alignment_scores:
Quality: 731.50 Length: 360
Ratio: 2.950 Gaps: 3
Percent Similarity: 68.889 Percent Identity: 43.611

alignment_block:
US-08-878-801-2 x US-08-371-930-28 ..
Align seg 1/1 to: US-08-371-930-28 from: 1 to: 2274

14 LeuThrGluAspGluLysAlaAlaArgValAspGlnGluIleAsnAr 30
||||| : : : : : : : : : : : : : : : : : : : :
36 CTGAGCGCAGAGGAGAGAGCGCCCTCGAGCGGAGCAAGCGGATTGAGAA 85
||||| : : : : : : : : : : : : : : : : : : : :
30 gIleLeuLeuGlnLysLysGlnAspArgGlyGluLeuLysLeuLeu 47
: ||| ||| : : : : : : : : : : : : : : : : : : : :
86 AAACCTCAAGAGAGATGCGATCAGCGCGCCCAAGACGCTGAATTTACTCC 135
||||| : : : : : : : : : : : : : : : : : : : :
47 euLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMetArg 63
||||| : : : : : : : : : : : : : : : : : : : :
136 TGTCTGGGGCTGGAGATCAGGAAAAAGACCATTGTGAAGCAGATGAG 185
||||| : : : : : : : : : : : : : : : : : : : :
64 IleIleHisGlyAlaGlyTyrSerGluGluArgLysGlyPheArgPr 80
||||| : : : : : : : : : : : : : : : : : : : :
186 ATCATCATGAAGATGGCTTCTCTGGGGAAGACGCTGAAGCAGTACAAGCC 235
||||| : : : : : : : : : : : : : : : : : : : :
80 oLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluAla 97
||||| : : : : : : : : : : : : : : : : : : : :
236 TGTGTCTACAGCAACACCATCCAGTCTCTGTGCGGCCCATTTGTCGGGCCA 285
||||| : : : : : : : : : : : : : : : : : : : :
97 etGluArgLeuGlnIleProPheSerArgProGluSerLysHisAla 113
||||| : : : : : : : : : : : : : : : : : : : :
286 TGGACACTTTGGCGTGGAGTGTGTGACAAGGAGGAGGAGGACGACTCC 335
||||| : : : : : : : : : : : : : : : : : : : :
114 SerLeuValMetSer.....GlnAspProTyrLysValTh 125
```

OM of: US-08-878-801-2 to: EST:*
Date: Sep 6, 2001 1:02 PM
out_format : pfs

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```

-MODEL=frame+ p2n_model -DEV=x1h
-Q/cgn2_1/USPO.spool/6004808/runat_06092001_110135_4462/app_query.fasta_1.437
-DB-EST -Qfmt=fastap -SUFFIX=p2n_rst -GAPOP=12.000 -GAPEXT=4.000
-NMATCH=0, 100 -LOOPEXT=0.000 -LOOPEXT=0.000 -CGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXT=7.000 -SPAIR=1.000 -YGAPEXT=0.500 -DELOP=6.000
-LIST=45 -LOCALIGN=200 -MATRIX=blosum62 -TRANS=human00.cd1
-ALIGN=15 -MODE=LOCAL -OUTFM=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=6004808_@cgn1_1.4084 -NCPU=6 -TCPU=3
-LOGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLXPY_WAIT
-THREADS=1

```

Search information block:

```

Search information block:
Query: US-08-878-801-2
Query length: 374
Database: EST.*
Database sequences: 10228115
Database length: 431459454
Search time (sec): 979.810000

```

score_list:

Sequence	Strd	Orig	zScore	EScore	Len	Document
gb_est29..AL532310	+	1087.50	2281.88	5.5e-118	909	AL532310 AL532310 LTI_NFL001..
gb_est80..BE915352	+	991.50	2077.51	1.3e-106	939	BE915352 601667635F1 NC1_CGAP..
gb_est29..AL541152	+	915.50	1939.72	1.4e-97	962	AL541152 AL541152 LTI_FLO02..PL1
gb_est29..AL552208	+	832.00	1715.72	1.5e-88	898	AL552208 LTI_FLO02..PL1
gb_est93..BF632767	+	800.00	1678.25	2.5e-88	898	BF632767 LTI_NFL006..PL1
gb_est46..AW366893	+	782.00	1678.25	2.5e-88	898	AW366893 LTI_FLO02..PL1
gb_est46..BF443684	+	778.00	1640.49	2.9e-82	470	BF443684 LTI-HT02156-201000-011
gb_est98..BG258764	+	777.50	1631.53	3.9e-81	1153	BG258764 602378022F1 NIH_MGC..S
gb_est87..BF101631	+	772.00	1620.33	3.9e-81	1153	BF101631 601753552F1 NC1_CGAP..S
gb_est79..BE788747	+	741.50	1545.85	5.5e-77	744	BE788747 601475843F1 NIH_MGC..S
gb_htc..AK004566	+	740.50	1529.48	4.5e-76	3261	AK004566 Mus musculus adult male
gb_est15..BF740136	+	730.00	1525.27	4.5e-76	768	BF740136 601595077F1 NIH_MGC..9
gb_est33..AA175435	+	730.00	1525.27	4.5e-76	768	AA175435 601595077F1 NIH_MGC..9
gb_est14..AA951538	+	726.50	1521.71	1.1e-75	890	AA951538 ms87g03..x1 Soares..send
gb_htc..AK009388	+	723.00	1510.14	5.4e-75	717	AK009388 LD32085..5prime LD Dros
gb_est90..BG248136	+	721.50	1497.06	2.9e-74	1658	BG248136 602400557F1 NIH_MGC..15
gb_est78..BE248136	+	713.50	1486.21	1.2e-73	985	BE248136 602400557F1 NIH_MGC..15
gb_est16..AII13347	+	703.00	1465.21	1.7e-72	881	AII13347 GH05204..5prime GH Dros
gb_est29..AL534767	+	668.00	1436.69	6.6e-71	663	AL534767 AL534767 LTI_FLO13..FBX
gb_est80..BE988440	+	667.00	1388.93	3.2e-68	837	BE988440 601512813F1 NIH_MGC..71
gb_est29..AL536202	+	650.00	1371.96	3.1e-67	1033	AL536202 AL536202 LTI_FLO13..FBX
gb_est18..AII325364	+	648.00	1370.67	3.1e-67	880	AII325364 601536022 LTI_FLO13..FBX
gb_est24..AW230435	+	642.00	1348.38	5.5e-66	878	AW230435 u063601..y1 Soares..mouse
gb_est29..AL505209	+	642.00	1338.94	1.8e-65	663	AL505209 u063601..y1 NC1_CGAP..Mous
gb_est15..AII569172	+	641.00	1334.25	3.4e-65	987	AII569172 AL569172 LTI_FLO02..PL1
gb_est28..AL536038	+	641.00	1335.41	2.1e-65	747	AL536038 6029406..x1 Soares..para
gb_est29..AL541652	+	636.50	1323.35	1.4e-64	923	AL541652 AL536038 LTI_FLO13..FBX
gb_est48..AW506450	+	619.00	1315.66	3.6e-64	943	AW506450 GH10384..5prime GH Dros
gb_est11..AA017829	+	618.00	1289.23	1.1e-62	712	AA017829 GH10384..5prime GH Dros
gb_est16..AII13889	+	602.00	1287.70	1.3e-62	677	AII13889 GH10384..5prime GH Dros
gb_est30..AII18196	+	602.00	1256.17	7.5e-61	549	AII18196 GH10384..5prime GH Dros
gb_est100..BG437180	+	600.00	1251.28	1.4e-60	831	BG437180 602490439F1 NIH_MGC..18
gb_est29..AL539321	+	600.50	1248.74	8.4e-60	915	AL539321 LTI_FLO13..FBX
gb_est29..AL538225	+	596.00	1237.40	8.4e-60	915	AL538225 LTI_FLO13..FBX
gb_est89..BF560715	+	592.00	1228.91	2.5e-59	914	BF560715 01-R-C1-kf-b-09-00-01..U
gb_est89..AL551940	+	590.00	1235.42	1.1e-59	367	AL551940 LTI_FLO02..PL1
gb_est100..BG395861	+	588.00	1220.09	7.7e-59	939	BG395861 602458647F1 NIH_MGC..16
gb_est100..BG395861	+	585.50	1216.21	1.3e-58	832	BG395861 602458647F1 NIH_MGC..16
gb_est29..AL565628	-	584.50	1211.81	2.2e-58	1009	AL565628 LTI_FLO13..FBX

```

gb_est46:AW429759      + 583.00 1220.94 6.9e-59      355      ! AW429759 68217 MARC lPIG_S
gb_est26:AI925377      + 582.00 1213.52 1.8e-58      666      ! AI925377 wp53d11.x1 NCI_CG
gb_est69:BE014065      + 583.00 1213.80 1.7e-58      543      ! BE014065 125642 MARC lPIG_
gb_est28:AL521216      - 581.00 1206.90 4.2e-58      814      ! AL521216 LTI_NFL001
gb_est29:AL548607      + 581.00 1206.04 4.7e-58      876      ! AL548607 AL548607 LTI_NFL001

seq_name: gb_est29:AL532310

seq_documentation_block:
LOCUS      AL532310
DEFINITION AL532310 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM014Y116 5
            prime, mRNA sequence.
ACCESSION  AL532310
VERSION    AL532310.1 GI:12795803
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 909)
AUTHORS   Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
            Location/Qualifiers
                1..909
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="CS0DM014Y116"
                    /clone_lib="LTI_NFL001_NBC4"
                    /sex="male"
                    /tissue_type="neuroblastoma cells"
                    /lab_host="DH10B"
                    /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
172 a 299 c 291 g 146 t 1 others
BASE COUNT
ORIGIN

```

alignment_scores:

```

Quality: 1087.50      Length: 220
Ratio: 5.012         Gaps: 1
Percent Similarity: 98.636      Percent Identity: 97.273

alignment_block:
US-08-878-801-2 x AL532310      . . .
Align seg 1/1 to: AL532310 from: 1 to: 909

1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
|||||
251 ATGGCCCGCTCGCTGACCTGGCGCTGCTGCCCTGGTGCCTGACGGAGA 300

17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeuG 34
|||||
301 TGAGAAGCCGCCGCCGCCGGTGGACCGAGGATCAACAGGATCCTCTTGG 350

34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuLeuGlyPro 50
|||||
351 AGCAGAGAAGACGAGGACCGCGGGAGCTGAAGCTGCTCTTTGGGCCCA 400

51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG1 67

```



```

239 TGCAGTGGCTGTGGAGGAGTCCCGCATCCGGCCTGCTATGAGCGTCGG 190
152 ArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLeuG1 168
|||||
189 CGGAATTCACCTGCTGCTGATTCAGCGCTGCTACTACCTGCTCCACCTGGA 140
|||||
168 uArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuArg 185
|||||
139 GCGCATCACCAGGAGGCTACGTCCTCCACAGCTCAGGACGTGCTCCGCA 90
|||||
185 erArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGlnLys 201
|||||
89 GCGCATGCCACCTACGTCATCAACGAGTACTGCTTCTCCGTCGCCGAAA 40
|||||
202 ThrAsnLeuArg 205
|||||
39 ACCAACCTGGCA 28

seq_name: gb_est46:AW368893

seq_documentation_block:
LOCUS AW368893 470 bp mRNA EST 04-FEB-2000
DEFINITION IL2-HT0199-241099-013-F04 HT0199 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW368893
VERSION AW368893.1 GI:6873543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 470)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL2&t2=IL2-HT0199-241099-013-F04&t3=1999-10-24&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 436.
Location/Qualifiers
1..470
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0199"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: Smal;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 121 a 138 c 121 g 90 t
ORIGIN

alignment_scores:
Quality: 782.00 Length: 151
Ratio: 5.213 Gaps: 0
Percent Similarity: 99.338 Percent Identity: 98.675

alignment_block:
US-08-878-801-2 x AW368893

```

Align seg 1/1 to: AW368893 from: 1 to: 470

```

185 SerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGlnLy 201
:::|||||
18 GCTCGAATGCCACCTGCGATGGCAGTCAACGAGTACTGCTTCTCCGTGCAGAA 67
|||||
201 sThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLysL 218
|||||
68 AACCAACCTTCGGATCGTGGACGTCGGGGCCAGAGTCAGAGCGTAAAG 117
|||||
218 ysTrpLleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaSer 234
|||||
118 AATGATCCATTTGTTGAGAACGTGATCGCCCTCATCTACCTAGCCTCA 167
|||||
235 LeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArgMe 251
|||||
168 CTGAGTGAATACGACCATGCTGGAGGAGAACCAACGAGGAGACCGCAT 217
|||||
251 tLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPheL 268
|||||
218 GAAGGAGAGCGCTCGCATTTGTTGGACTATCTCTGGAACCTACCTGGTCA 267
|||||
268 ysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluGlu 284
|||||
268 AAGCACATCGTCTCTCTTCTCAACAAACCGACATCTCTGGAGGAG 317
|||||
285 LysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGlyPr 301
|||||
318 AAAATCCCACTCCCTCCCTGGCTGCTATCTTCCCGAGTTTCCAGGGCCC 367
|||||
301 OLYSGlnAspAlaGluAlaAlaLysArgPheIleLeuAspMetTyrThrA 318
|||||
368 TAACGAGATGCTGAGGAGCAGAGAGTTTCATCTCGACATGTACAGCA 417
|||||
318 rgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyAla 334
|||||
418 GGATGTACACCGTGTGCGTGGCGTCCCGAGGCGCAGCAAGAGGGCGCA 467
|||||
335 Arg 335
|||
468 CGA 470

```

seq_name: gb_est87:BF443684

seq_documentation_block:
LOCUS BF443684 489 bp mRNA EST 01-DEC-2000
DEFINITION 261269 MARC 2F1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF443684
VERSION BF443684.1 GI:11503776
KEYWORDS EST.
SOURCE pig.

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE

1 (bases 1 to 489)

AUTHORS

Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,

Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.

and Keefe,J.W.

Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt.trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACACGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGCG
 Plate: 94 row: L column: 13
 Seq primer: ATTTAGGTGACACTATAG.

FEATURES

Source
 1. 489
 Location/Qualifiers
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 2PIC"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
 Library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 117 a 145 c 132 g 93 t 2 others
 ORIGIN

alignment_scores:
 Quality: 778.00 Length: 162
 Ratio: 4.955 Gaps: 1
 Percent Similarity: 96.914 Percent Identity: 93.210

alignment_block:

US-08-878-801-2 x BF443684 ..

Align seg 1/1 to: BF443684 from: 1 to: 489

205 ArgileValAspValGlyGlyGlnLysSerGluArgLysLysThrIleHi 221
 5 AGGATCGTGATGTTGGGGCCATGAGTCAGACGCAAGAGTGGATCCA 54
 221 sCysPheGluAsnValIleAlaLeuIleThrLeuAlaSerLeuSerGluT 238
 55 CTGCTTCTCAATGTCATCGCCCTCATCTACCTAGCTCGCTGAGTGAAT 104
 238 rAspGlnCysLeuGluGluAsnGlnCysAspArgMetLysGluSer 254
 105 ATGACCACTGCTCGAGGAGAAACACCAACAGAGAACCAATGAAGGAAGC 154
 255 LeuAlaLeuPheGlyThrIleLeuGluLeuProThrPheLysSerThrSe 271
 155 CTGCGCCCTGTCGGACCATCTTGGAACTGCGCTGTTCAAAAGCACTTC 204
 271 rValIleLeuPheLeuAsnLysThrAspIleLeuGluLysIleProT 288
 205 TGTATCTCTCTCTCACAAACACGACATCTCTGGAGGAGAGATCCCCA 254
 288 hrSerHisLeuAlaThrTyrPheProSerPheGlnGlyProLysGlnAsp 304
 255 CTTCCACCTCTGGTACTTCTCTCCAGTTCCAGGCGCCGGAAGAGGAC 304
 305 AlaGluAlaAlaLysArgPheIleLeuAspMetTyrThrArgMetTyrTh 321
 305 GCGAGGACGACCAAGAGGTTCATCTGGACATGTACACCAAGATGATCGC 354
 321 rGlyCysValAspGlyProGluCysLysLysLysGlyAlaArgSerArgA 338
 355 CGGCTCGTGATGGCGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 404
 338 rgLeuPheSerHisTyrThrCysAlaThrAspThrGlnAsnIleArgLy 354
 405 GNCCTTTTACGCACATACAGTGTGCGACGACACACACATATCCGCA 454
 354 sValPheLysAspValArgAspSerValLeuAla 365
 455 GGTCTTCAGGAGCTGCGGACGCTCGGTCTCTGGCC 488

seq_name: gb_est98:BG258764

seq_documentation_block: 1153 bp mRNA EST 13-FEB-2001
 LOCUS BG258764
 DEFINITION 602378022P1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4509052 5',
 mRNA sequence.

ACCESSION

BG258764
 BG258764.1 GI:12768580

EST.

human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1153)

REFERENCE

NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10388 row: k column: 05

High quality sequence stop: 720.

FEATURES

Source

1. 1153

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4509052"

/clone_lib="NIH_MGC_52"

/tissue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

225 a 405 c 369 g 154 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 777.50 Length: 192

Ratio: 4.418 Gaps: 6

Percent Similarity: 91.667 Percent Identity: 84.896

alignment_block:

US-08-878-801-2 x BG258764 ..

Align seg 1/1 to: BG258764 from: 1 to: 1153

1 MetAlaArgSerLeuThrTrpArgCysCysPrcTrpCysLeuThrGluAs 17
 308 ATGGCCCGCTGCTGACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
 17 pGluLysAlaAlaArgValaspGlnGluIleAsnArgIleLeuLeuG 34
 358 TGAGAGGCGCCCGCCCGGGTGGACAGAGAGNCAACAGGATCTCTTGG 407
 34 lucInLysLysGlnAspArgGlyCysLeuLysIleLeuLeuLeuGlypro 50
 408 AGCAG 457
 458 GGCAG 507
 51 GlyCysSerGlyLysSerThrPheIleLysGlnMetArgIleLeuHisG 67
 67 yAlaGlyTyrSerGluCysLysGlyPheArgProLeuValTyrG 84
 508 CGCCGGCTACTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 557
 84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
 558 AGAACATCTTCTGCTCCATCGGGCCCATGATCCAGCCCTGGAGCGGCTG 607


```

1162 GATGTCATCAATAAAACAACCTTAAGGAATGTGGCCT 1200
||||| ..... ||| ||| ..... |||
seq_name: gb_est78:BE740136
seq_documentation_block:
LOCUS BE740136 768 bp mRNA EST 15-SEP-2000
DEFINITION 601595077F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949146 5',
mRNA sequence.
ACCESSION BE740136
VERSION BE740136.1 GI:10154128
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 768)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC813 row: 1 column: 19
High quality sequence stop: 735.
FEATURES
Location/Qualifiers
1..768
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949146"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="PH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site: 1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 171 a 241 c 236 g 130 t
ORIGIN

```

```

alignment_scores:
Quality: 730.50 Length: 258
Ratio: 3.462 Gaps: 5
Percent Similarity: 81.783 Percent Identity: 58.527

```

```
alignment_block:
```

```
us-08-878-801-2 x BE740136 ..
```

```
Align seg 1/1 to: BE740136 from: 1 to: 768
```

```

36 LysLysGlnAspArgGlyGluLeuLysLysLeuLeuLeuGlyProGlyG1 52
|||||:||||| ||| ||||||||| ||||||||| |||||||
11 AAGCGGACCCCGCGCGGAGTCAAGTCTGCTCGCGCACGGCGA 60
52 uSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisGlyAlaG 69
61 GAGCGGGAAGAGCAGCTTCATCAAGCAGATGCGCATCATCCACGGCGG 110
69 lYrSerGluGluArgLysGlyPheArgProLeuValTyrGlnAsn 85
|||||:||||| ..... ||| ||||||||| |||||||||
111 GCTACTCGGAGGAGGAGGCGGCTTCCACCAAGCTCGTCTACCAAGAAC 160
86 IlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeuGlnI1 102

```

```

|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
161 ATCTTTCACCGCCATGAGGCCATGATCCGGGCCATGGACAGCTCAAGAT 210
102 eProPheSerArgProGluSerLysHisAlaSerLeuAlaMetSerG 119
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
211 CCTCTACAAGTACGAGCAGACAAGGCCAATGC3CTCCTGATCCGGGAGG 260
119 lAsAsProTyrLysValThrPheGluLysArgTyrAlaAlaMet 135
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
261 TGGACGTGGAGAAGGTGACCACTTCGAGCATCAGTACGTGAGTCCATC 310
136 GlnTyrLeuTyrArgAspAlaGlyIleArgAlaLysTyrGluArgArg 152
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
311 AAGACCCGTGGAGAGACCCGGCATCCAGGAATGCTACGACCCGAGGCG 360
152 gGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLeuGlu 169
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
361 CGAGTACCAAGCTCTCCGACTCTGCAAGTACTA:CTGACCCGACGTTGAC 410
169 rGileThrGluGluGlyTyrValProThrAlaGlnAspValLeuArgSer 185
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
411 GCATCGCCACCTTGGGCTACCTGCCACCCAGCAGGAGCTGCTGGGGTC 460
186 ArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGlnLysTh 202
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
461 CGCGTCCGACACCGGATCATCGAGTACCT:TCGACCTGGAGAACAT 510
202 rAsnLeuArgIleValAspValGlyGlnLysSerGluArgLysLysT 219
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
511 CATCTTCGGGATGTGGATGTGGGGGCCAGCG:TCGAGCGGAGGAGT 560
219 rPileHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaSerLeu 235
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
561 GGATCCACTGCTTTCAGAACCTGACATCCATCA:GTTTCTGTCGCCCTC 610
236 SerGluTyrAspGlnCysLeuGluGluAsnAsnGln:GluAsnArgMetL 252
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
611 AGCAATACGACCAAGTCTGCTGGAGTCCGAC:AAACGAGAACCGGATGG 660
252 ysGluSerLeuAlaLeuPheGlyThrIleLeuGlu:.....LeuProTrp 266
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
661 AGGAGAGCAAGCCCTGTTCCGACCATCATCA:TAACCTGGTTCCG... 707
267 PheLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuG1 283
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
708 .....AATCTCTCCGTCATCTCTTC...TCACAAGAAGGAGCTGCTGGA 748
283 uGluLysIleProThrSerHis 290
|||||:||||| ..... ||| ||||||||| ||||||||| |||||||
749 GGACAGA...TCTGTTTCGCAC 767

```

```
seq_name: gb_est15:AI040021
```

```
seq_documentation_block:
```

```

LOCUS AI040021 890 bp mRNA EST 24-SEP-1998
DEFINITION ox97g03.x1 Soares_senescent_fibroblasts_NbHSF_Homo_sapiens_cDNA
clone IMAGE:1664308 3', similar to gb:M69013 GUANINE
NUCLEOTIDE-BINDING PROTEIN G(Y), ALPHA SUBUNIT (HUMAN);, mRNA
sequence.
ACCESSION AI040021
VERSION AI040021.1 GI:3279215
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

```

This clone is available royalty-free through LNL : contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1033 Std Error: 0.00
 Seq primer: -40m13 fwd, R from Amersham
 High quality sequence stop: 497.

FEATURES

source

```

1. .890
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1664308"
/clone_lib="Soares_senescent_fibroblasts_NbHSP"
/tissue_type="senescent_fibroblast"
/lab_host="Dh10B (ampicillin resistant)"
/notes="Vector: p773D (Pharmacia) with a modified
polylinker V.TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(dt)
primer [5',
TGTACCACTCTCAAGTGGGAGCGCGCATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT      199 a      263 g      153 t      3 others
ORIGIN

```

```

alignment_scores:
  Quality: 730.00      Length: 284
  Ratio: 3.333        Gaps: 2
  Percent similarity: 77.113 Percent Identity: 53.873

```

alignment_block:

US-08-878-801-2 x AI040021 ..

Align seg 1/1 to: AI040021 from: 1 to: 890

```

13 CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluLeuAs 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 TGCCTGAGCATGAGTGAAGGAGTCAACCGGATCAACCGCGAGATCGA 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 nArgIleuLeuGluGlnLysLysGlnAspArgGlyGluLeuLeuL 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
82 GAAGCAGCTCGCGCGGACAAAGCGCGAGCGCGCGGAGCTCAAGCTGC 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 euLeuLeuGlyProGlySerGlyLysSerThrPheIleLysGlnMet 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 TGCTGCTGCGGACGCGGAGAGCGGGAAGACGACGTTTCATCAAGCAGATG 181
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 ArgIleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 CGCATCATCCACGCGCGGCTACTCGGAGGAGGACAAAGCGCGCTTCAC 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 CAAGCTCGCTACCAAGACATCTTCAACCGCATCGAGGCCATGATCGGG 281
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHis 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 CCATGGAGACGCTCAAGATCTCTCAACAGTACGAGCAAGAACAGCCCAAT 331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLy 129
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 GCGCTCTGATCCGGAGGTGACCTGGAGAGGTGACCACTTCGAGCA 381
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 sArgTyrAlaAlaIaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 TCAGTACGTACGTGCGCATCAAGACCTGTGGGAGGACCGCGCATCCAGG 431
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 AATGCTACGACCGCGAGCGGAGTACCACTCTCCGACTCTCCCAAGTAC 481

```

```

163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 TACTGACCGACGTTGACCGCATCGNCACCTTGGGCTACCTGCCACCCA 531
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
532 GCAGGACGTGCTCGGGTCCGCGTCCGCCACCGCGCATCATCGAGTACC 581
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
582 CTTTCGACCTGGAGACATCATCTTNCGGATGCTGGATGTGGGGGCCCA 631
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 LysSerGluArgLysTyrPheHisCysPheGluAsnValIleAlaLe 229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
632 CGGTCGCGAGCGAGGAGTGCCTACTGCTNTGAGAACCTGACATCCAT 681
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
682 CATGTTTCTCGTCCCTCAGCGAATACGACCAAGTCTCTGTTGGGAGTCG 731
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
732 ACCAGGAGACCGGATGGAGGAGAGCAAGCCCTGTC.CGGGACATCATC 780
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 ACCTACCCCTGGTCCAGACTCTTCGTCATCTCTTCTTACCAAGGAAG 830
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
279 rAspIleLeuGluGlyLysIleProThrSerHisLeuAlaThrTyrPheP 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
831 ACCT...GCTGGAGACAAGTACTGACTTGC...ACTGGTGGCTACTTCC 874
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 ro 296
||
875 CC 876

```

seq_name: gb_est3:AA175435

```

seq_documentation_block:
LOCUS      AA175435      505 bp      mRNA      EST      16-FEB-1997
DEFINITION      ms87c01.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:618528
                5', similar to gb:M80632 Mouse G protein alpha subunit (MOUSE)).
                mRNA sequence.

```

```

ACCESSION      AA175435
VERSION        AA175435.1
GI:1756593

```

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL : contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:379352
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 451.

FEATURES

source

Location/Qualifiers

1..505

Thu Sep 6 15:16:55 2001

```

204 AGGATAATTGAGTATCCCTTTGATTAGAGAAATCAGATTTAGAATGG 253
207 aAspValGlyGlyGlnLysSerGluArgLysTrpIleHisCysPhe 223
254 TAGACGTCGGTGTGTCAGCGCAGAGAGAAAGTGGATTTCATTCCTTT 303
224 GluAsnValIleAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspG 240
304 GAGAAATGTGACATCAATATATATTTTGGTAGCGCTATCGGAGTACGATCA 353
240 nCysLeuGluGluAsnGlnGluAsnArgMetLysGluSerLeuAlaL 257
354 AATCTTGTGTAATCGAATGAGAAATCGAATGGAGGAATCAAAGCCT 403
257 euPheGlyThrIleLeuGluLeuProTyrPheLysSerThrSerValIle 273
404 TATTCGTACTATTAATACATACCTTGGTTTCAAAATTCGTCAGTTATT 453
274 LeuPheLeuAsnLysThrAspIleLeuGluGluLysIleProThrSerHi 290
454 CTTTTCTCTGAATAAGAAGGACTTGTGGAAGAGAGAAATAATGTATTCGCA 503
290 sLeuAlaThrTyrPheProSerPheGlnGlyProLysGlnAspAlaGluA 307
504 TTTGGTAGACTATTTTCTGAATACGATGGTCCFCACGAGATGCCAATAA 553
307 laAlaLysArgPheIleLeuAspMetTyrThrArgMetTyrThrGlyCys 323
554 CGGCCCGAGAGTTTATCTCGGAATGTTGTGATTTTA..... 591
324 ValAspGlyProGluGlySerLysLysGlyAlaArgSerArgArgLeuPh 340
592 .....AATCCAGATTCGGAANA.....ATTACTA 617
340 eSerHisTyrThrCysAlaThrAspThrGlnAsnIleArgLysValPheL 357
618 TTCTCATTTTCACGTGTGCTACAGATACGGAATAATATAAGTTTGTGTTG 667
357 ysAspValArgAspSerValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
668 CAGCTGTAAAGGACACAATCTGCAATCGAACCTTAAGGAATATAATTTG 717

```